



Parcellisation du manteau cortical à partir du réseau de connectivité anatomique cartographié par imagerie de diffusion

Pauline Roca

► To cite this version:

Pauline Roca. Parcellisation du manteau cortical à partir du réseau de connectivité anatomique cartographié par imagerie de diffusion. Autre [cond-mat.other]. Université Paris Sud - Paris XI, 2011. Français. NNT : 2011PA112225 . tel-00652673

HAL Id: tel-00652673

<https://theses.hal.science/tel-00652673>

Submitted on 16 Dec 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ PARIS SUD XI, FACULTÉ DES SCIENCES
D'ORSAY

École Doctorale STITS
(Sciences et Technologies de l'Information et de la Communication)

THÈSE

en vue d'obtenir le grade de

Docteur, spécialité « Traitement du Signal »

présentée et soutenue par

Pauline ROCA

PARCELLISATION DU MANTEAU CORTICAL À PARTIR DE LA CONNECTIVITÉ ANATOMIQUE CARTOGRAPHIÉE PAR IMAGERIE DE DIFFUSION

Date prévue de soutenance, 31 Septembre 2011

Composition du jury :

<i>Rapporteurs</i>	Dr ISABELLE BLOCH	CNRS, Télécom ParisTech, France
	Pr JEAN-PHILIPPE THIRAN	Ecole Polytechnique de Lausanne, Suisse
<i>Examineurs</i>	Dr IRÈNE BUVAT	CNRS, Université Paris Sud XI, France
	Dr OLIVIER COULON	CNRS, UMR 6168, France
	Pr STÉPHANE LEHERICY	CENIR, ICM, France
	Dr DENIS RIVIÈRE	LNAO, Neurospin, CEA, France
<i>Directeur de thèse</i>	Dr JEAN-FRANÇOIS MANGIN	LNAO, Neurospin, CEA, France
<i>Membre Invité</i>	Dr CYRIL POUPON	LRMN, Neurospin, CEA, France

A mes parents.

REMERCIEMENTS

Quoique peuvent en dire les thésards en déprime de fin de thèse ;-), effectuer sa thèse à Neurospin est vraiment une chance.

Je voudrais tout d'abord remercier mon directeur de thèse, Jean-François Mangin, pour ces conseils, sa patience, son écoute, et la liberté et la confiance qu'il m'a accordées pendant cette thèse. Je le remercie également pour ses précieux conseils sur mon avenir professionnel post-thèse et sa disponibilité pour parler de ce sujet.

Je remercie sincèrement Isabelle Bloch et Jean-Philippe Thiran d'avoir accepté d'être les rapporteurs lors de ma soutenance de thèse, ainsi que Irène Buvat, Olivier Coulon, et Stéphane Lehericy qui ont accepté de participer à mon jury.

Je remercie tout particulièrement Denis Rivière, qui m'a été d'une aide essentielle dans la mise en place des outils de visualisation, compte tenu de mes faibles compétences initiales en informatique. Je le remercie également pour tous ces covoiturages du matin avec le compère Grégory, pour la musique, le théâtre rouge, sans oublier toutes les discussions et remontages de moral qui m'ont beaucoup aidé. Pour tout ça et d'autres choses que j'oublie sûrement, merci milles fois.

Je remercie également Cyril Poupon, pour sa gentillesse, ces compétences précieuses en imagerie de diffusion et le temps qu'il m'a consacré malgré ces nombreuses occupations. Je le remercie également d'avoir acquis la base de données, indispensable pour ma thèse ;-).

Je remercie aussi les thésards avec qui j'ai eu la chance de collaborer, Pamela, Linda, et, Alan, pour leur gentillesse, leur générosité et le plaisir que j'ai eu à travailler avec eux.

Pour m'avoir supporté au quotidien, je remercie les collègues de Neurospin qui ont partagé mon bureau et ensuite l'open-space (et bien d'autres choses...). Merci à Dominique, Eric, Clara, Edith, Cécilia et Valdis pour leur soutien, leurs conseils, et leur bonne humeur.

Je n'oublie pas les autres thésards et permanents également : merci à Laurent (pour son accueil londonien et ses précieux conseils pré-entretiens), Merlin, Solveig, Soizic, Grégory, Benjamin, Yannick, Vincent, Thomas, Fabian, Gaël, Anne-Laure, Véronique, Irina, Jimmy, François, Dimitri, Edouard, Isabelle, Yann, et tous les autres...

Sans oublier Maryline et Denis (Fournier) qui m'ont bien aidé pour la préparation pratique de ma soutenance.

Merci à vous tous d'avoir contribué à faire du laboratoire un lieu où il fait bon travailler.

Je remercie de tout mon coeur mes amis, avec une mention spéciale à Katline avec qui j'ai pu partager malgré la distance beaucoup de choses, merci au chevalier Gambette notamment ;-). Merci encore.

Je n'aurais pas accompli ce travail sans le soutien de ma famille. Merci à mes parents et à mes soeurs, Maryse, Louis, Laëtitia et Marie-Laure ainsi qu'à Dolorès, Victor, Clément, Vincent, Jeanne, Pierre, Benjamin, Lilya, Anne, Didier, Arthur et Sarah.

Je conclurai en remerciant de tout coeur Simon, mon homme !

Paris, le 16 décembre 2011.

TABLE DES MATIÈRES

TABLE DES MATIÈRES	vii
INTRODUCTION	1
I Le connectome cérébral humain	5
1 CONTEXTE NEUROSCIENTIFIQUE	7
1.1 GÉNÉRALITÉS	9
1.2 PLUSIEURS DÉFINITIONS DES NOEUDS DU CONNECTOME	18
1.3 PLUSIEURS DÉFINITIONS DES CONNEXIONS DU CONNECTOME .	26
1.4 CONNECTOME ET VARIABILITÉ INTER-INDIVIDUELLE	28
1.5 CONCLUSION	34
2 QUELS OUTILS POUR ACQUÉRIR QUEL CONNECTOME ET COMMENT ?	35
2.1 UNE ANALYSE STRUCTURELLE SURFACIQUE	37
2.2 ACQUISITION DES IMAGES IRM	39
2.3 PRÉTRAITEMENTS DES IMAGES IRM	44
2.4 RÉCAPITULATIF DES DONNÉES PRÉTRAITÉES DISPONIBLES . . .	55
2.5 CONCLUSION	56
II Contributions	59
3 OUTILS SPÉCIFIQUES AU CONNECTOME ANATOMIQUE SURFACIQUE CORTICAL	61
3.1 INTRODUCTION	63
3.2 PARCELLISATIONS ALÉATOIRES DU CORTEX	65
3.3 MATRICE DE CONNECTIVITÉ CORTICALE	67
3.4 VISUALISATION DE LA MATRICE DE CONNECTIVITÉ	69
3.5 CONCLUSION	72
4 CONSTRUCTION DE CONNECTOMES SIMULÉS	75
4.1 INTRODUCTION	77
4.2 CONSTRUCTION DES DONNÉES SIMULÉES	78
4.3 DONNÉES UTILISÉES POUR LA VALIDATION	86
4.4 CONCLUSION	90
5 RÉDUCTION DES PROFILS DE CONNECTIVITÉ	91
5.1 INTRODUCTION	93
5.2 IDENTIFICATION AUTOMATIQUE DE RÉGIONS D'INTÉRÊT	95

5.3	ANALYSE EN COMPOSANTES PRINCIPALES	111
5.4	DIFFÉRENTES MESURES DE DISTANCE ENTRE PROFILS	122
5.5	CONCLUSION	128
6	PARCELLISATION ITÉRATIVE INTRA-SUJET DU CORTEX ENTIER	131
6.1	DONNÉES ET HYPOTHÈSES	133
6.2	MÉTHODE	135
6.3	ESSAIS SUR LES DONNÉES SIMULÉES	138
6.4	ESSAIS SUR LES DONNÉES RÉELLES	141
6.5	CONCLUSION	144
7	PARCELLISATION DE GROUPE DU CORTEX	145
7.1	PARCELLISATION AVEC L'IDENTIFICATION DE RÉGIONS D'INTÉRÊT	147
7.2	COMPARAISON DE DIFFÉRENTES MÉTHODES	158
7.3	CONCLUSION	171
	CONCLUSION GÉNÉRALE	173
	Annexes	183
A	REVUE DES PARCELLISATIONS CORTICALES À PARTIR DE LA CONNECTIVITÉ ANATOMIQUE	183
B	ÉLÉMENTS DE THÉORIE DES GRAPHS	187
	Références	191
	BIBLIOGRAPHIE	191
	PUBLICATIONS	205
	NOTATIONS ET ABRÉVIATIONS	207
	LISTE DES FIGURES	208
	RÉSUMÉ	212
	ABSTRACT	212

INTRODUCTION

La parcellisation du cerveau humain en aires fonctionnelles est un problème complexe mais majeur pour la compréhension du fonctionnement du cerveau et pourrait avoir des applications médicales importantes en neurochirurgie par exemple pour mieux identifier les zones fonctionnelles à sauvegarder. Cet objectif va de pair avec l'idée de construire le connectome cérébral humain, qui n'est autre que le réseau de ses connexions. Pour définir un tel réseau, il faut en effet définir les éléments de base : c'est-à-dire avoir un découpage du cerveau en régions. Il existe de multiples manières et critères pour identifier ces régions et à ce jour il n'y a pas de parcellisation universelle du cortex.

Dans cette thèse nous étudierons la possibilité d'effectuer cette parcellisation en fonction des données de connectivité anatomique, issues de l'imagerie par résonance magnétique de diffusion, qui est une technique d'acquisition permettant de reconstruire les faisceaux de neurones cérébraux de manière non invasive. Nous nous placerons dans un cadre surfacique en étudiant seulement la surface corticale et les connexions anatomiques sous-jacentes.

L'étude de ce type de connectome est relativement nouvelle et en premier lieu se pose la question du choix et de l'existence d'outils adaptés dans ce contexte. L'approche surfacique réduit la dimension des données à traiter, comparée à l'approche volumique, mais une réduction de dimension des données reste nécessaire.

Un des challenges les plus importants est la prise en compte de la variabilité interindividuelle des données. Comment la gérer, la modéliser, et la visualiser dans son ensemble dans le but de trouver des régions similaires pour un groupe de sujets simultanément ? Ainsi, avant de déterminer le nombre de régions que l'on peut identifier, l'un des enjeux est déjà de trouver une base de décomposition des données commune à l'ensemble des sujets sur lequel une parcellisation sera ensuite effectuée.

Enfin, un autre point important est l'identification des régions qui vont jouer un rôle dans la parcellisation d'une zone du cortex.

Contributions Nous proposons dans cette thèse un ensemble de nouveaux outils pour construire et visualiser le connectome cérébral humain, dans un cadre surfacique et à partir des données de connectivité anatomique reconstruites par IRM. Nous fournissons de plus une méthode pour simuler des connectomes pour un groupe de sujets en contrôlant la variabilité inter-individuelle des données. A partir de ces outils nous présenterons des méthodes de réduction de dimension des données de connectivité, que nous appliquerons pour parcelliser le cortex entier de quelques sujets. Nous proposons aussi une nouvelle manière de décom-

poser les données de connectivité au niveau d'un groupe de sujets en tenant compte de la variabilité inter-individuelle. Cette méthode sera testée et comparée à d'autres méthodes sur les données simulées et les données réelles.

Organisation de la thèse Ma thèse est divisée en deux parties séparant son cadre des contributions que j'ai réalisées. En voici un bref aperçu.

PARTIE I. LE CONNECTOME CÉRÉBRAL HUMAIN

La première partie de cette thèse, intitulée **Partie I. Le connectome cérébral humain** introduira le contexte de cette thèse, le choix du cadre et des données utilisées ainsi que les prétraitements effectués à partir de l'existant.

Contexte neuroscientifique

Le **Chapitre 1** posera le contexte neuroscientifique de cette thèse en présentant le connectome cérébral humain et les multiples façons de le construire et de le définir. Différentes échelles spatiales peuvent être envisagées, ainsi que différentes techniques de reconstruction allant de l'imagerie par résonance magnétique à la magnéto-encéphalographie. A cette multiplicité des connectomes se rajoutent les difficultés liées à la variabilité inter-individuelle que nous étudierons dans la fin de ce chapitre.

Quels outils pour acquérir quel connectome et comment ?

Nous préciserons dans le **Chapitre 2** le cadre de cette thèse, i.e. le type de connectome choisi parmi les définitions présentées dans le chapitre 1. Nous présenterons ainsi les choix techniques que nous avons faits, ainsi que les principes de l'acquisition des données nécessaires à la construction de ce connectome. Nous présenterons enfin les prétraitements que nous avons effectués pour modéliser la variabilité inter-individuelle et reconstruire les faisceaux de fibres à partir des outils disponibles dans la communauté scientifique.

PARTIE II. CONTRIBUTIONS

La deuxième partie intitulée **Partie II. Contributions** présentera les contributions apportées par cette thèse avec cinq chapitres suivant le cheminement que j'ai eu pendant mes années de thèse.

Outils spécifiques au connectome anatomique surfacique cortical

Dans le **Chapitre 3** nous présenterons les outils spécifiques que nous avons été amenés à développer pour construire et visualiser le type de connectome étudié dans le cadre de cette thèse. Nous proposons notamment une méthode de visualisation originale du connectome d'un sujet, mettant en évidence la répartition spatiale des connexions.

Construction de connectomes simulés

Dans le [Chapitre 4](#) nous proposerons une façon de construire des connectomes simulés pour un groupe de sujets, en modélisant la variabilité inter-individuelle des données ainsi qu'en contrôlant la quantité de connexions disponibles. Cette étape sera utilisée dans la suite de la thèse pour tester et comparer les méthodes de parcellisation du cortex.

Réduction des profils de connectivité

Dans le [Chapitre 5](#) nous présenterons une nouvelle manière de résumer et regrouper les données de connectivité sans utiliser de segmentation a priori du cortex et en tenant compte de la répartition spatiale des connexions, cette réduction de dimension pouvant être effectuée aux niveaux intra-sujet et inter-sujets. Nous essaierons de plus une technique plus classique de réduction de dimension et aborderons différentes mesures de distances possibles pour comparer les profils de connectivité.

Parcellisation itérative intra-sujet du cortex entier

Nous appliquerons ensuite dans le [Chapitre 6](#) la méthode de réduction de dimension précédente pour parcelliser le cortex entier de plusieurs sujets indépendamment en utilisant une procédure itérative et toujours sans utiliser une segmentation a priori. Cette démarche intra-sujet sera laissée de côté en faveur d'une parcellisation au niveau d'un groupe de sujets entier que nous effectuerons dans le chapitre suivant.

Parcellisation de groupe du cortex

Dans le [Chapitre 7](#) nous appliquerons les méthodes de réduction des données de connectivité pour identifier pour un groupe de sujets les données pertinentes à garder en vue d'un clustering inter-sujets. Nous comparerons différentes méthodes de réduction de dimension, avec ou sans utiliser une segmentation a priori sur les données simulées au Chapitre 4. Nous pourrons grâce à cette étude évaluer la pertinence et le cadre d'utilisation des différentes méthodes.

Dans la [conclusion](#) nous récapitulerons les parties précédentes ainsi que les contributions associées et nous évoquerons les perspectives des travaux effectués pendant cette thèse.

Première partie

**Le connectome cérébral
humain**

CONTEXTE NEUROSCIENTIFIQUE

1

L'ÉTUDE du connectome cérébral suscite actuellement un engouement dans la communauté neuroscientifique (Bressler et Menon 2010). De nombreuses applications tant scientifiques que médicales y sont associées. La théorie des graphes fournit dans ce contexte des outils pour étudier les propriétés de ce connectome.

Notons cependant la complexité de ce connectome qui commence par la multiplicité des définitions possibles. Différents choix sont envisageables en ce qui concerne les éléments de base et l'échelle spatiale associée ainsi que concernant la nature des connexions étudiées.

Nous allons dans les deux premières parties de ce chapitre faire un bref inventaire de toutes ces possibilités, en commençant par les différentes échelles spatiales envisagées ainsi que les applications et propriétés s'y rapportant. Nous aborderons ensuite brièvement les différentes définitions possibles des éléments de base et les connexions associées. Pour finir, nous parlerons de la variabilité inter-individuelle de ce connectome.

Ces éclaircissements sont nécessaires pour comprendre le contexte de cette thèse ainsi que son cadre c'est-à-dire le type de connectome étudié et les outils associés, cadre qui sera abordé dans la conclusion de ce chapitre ainsi que dans le chapitre suivant.

SOMMAIRE

2.1	UNE ANALYSE STRUCTURELLE SURFACIQUE	37
2.1.1	Avantages de l'approche surfacique	37
2.1.2	Une base de données adaptée à l'étude structurelle	38
2.2	ACQUISITION DES IMAGES IRM	39
2.2.1	Le phénomène de Résonance Magnétique Nucléaire	39
2.2.2	Encodage spatial des images	40
2.2.3	Imagerie anatomique : Contrastes en T_1 et T_2	41
2.2.4	IRM de diffusion : pour accéder aux tracts	41
2.2.4.1	Phénomène de diffusion	41
2.2.4.2	Imagerie et Séquences d'acquisition	42
2.3	PRÉTRAITEMENTS DES IMAGES IRM	44
2.3.1	Les logiciels utilisés	44
2.3.2	Prétraitements anatomiques	45
2.3.2.1	Extraction de la surface corticale : exemple de Brainvisa	45
2.3.2.2	Modélisation de la variabilité interindividuelle corticale de Freesurfer	46
2.3.2.3	Définition de régions corticales d'intérêt	46
2.3.3	Prétraitements des images de diffusion	49
2.3.3.1	Données locales de diffusion	49
2.3.3.2	Modèles locaux et autres techniques	50
2.3.3.3	Tractographie	53
2.4	RÉCAPITULATIF DES DONNÉES PRÉTRAITÉES DISPONIBLES	55
2.4.1	Premiers traitements	55
2.4.2	Seconds prétraitements	55
2.5	CONCLUSION	56

1.1 GÉNÉRALITÉS

1.1.1 Définition

Le **connectome cérébral humain** peut être défini comme le réseau des connexions cérébrales. On peut le modéliser par un graphe où chaque **noeud** sera une zone du cerveau et une relation entre deux noeuds les **connexions** les reliant (Sporns et al. 2005). Il existe de multiples façons de construire le connectome qui sont le reflet des différentes échelles spatiales choisies ainsi que des techniques de reconstruction des connexions associées dont la nature est aussi multiple. La figure 1.1 en donne un aperçu.

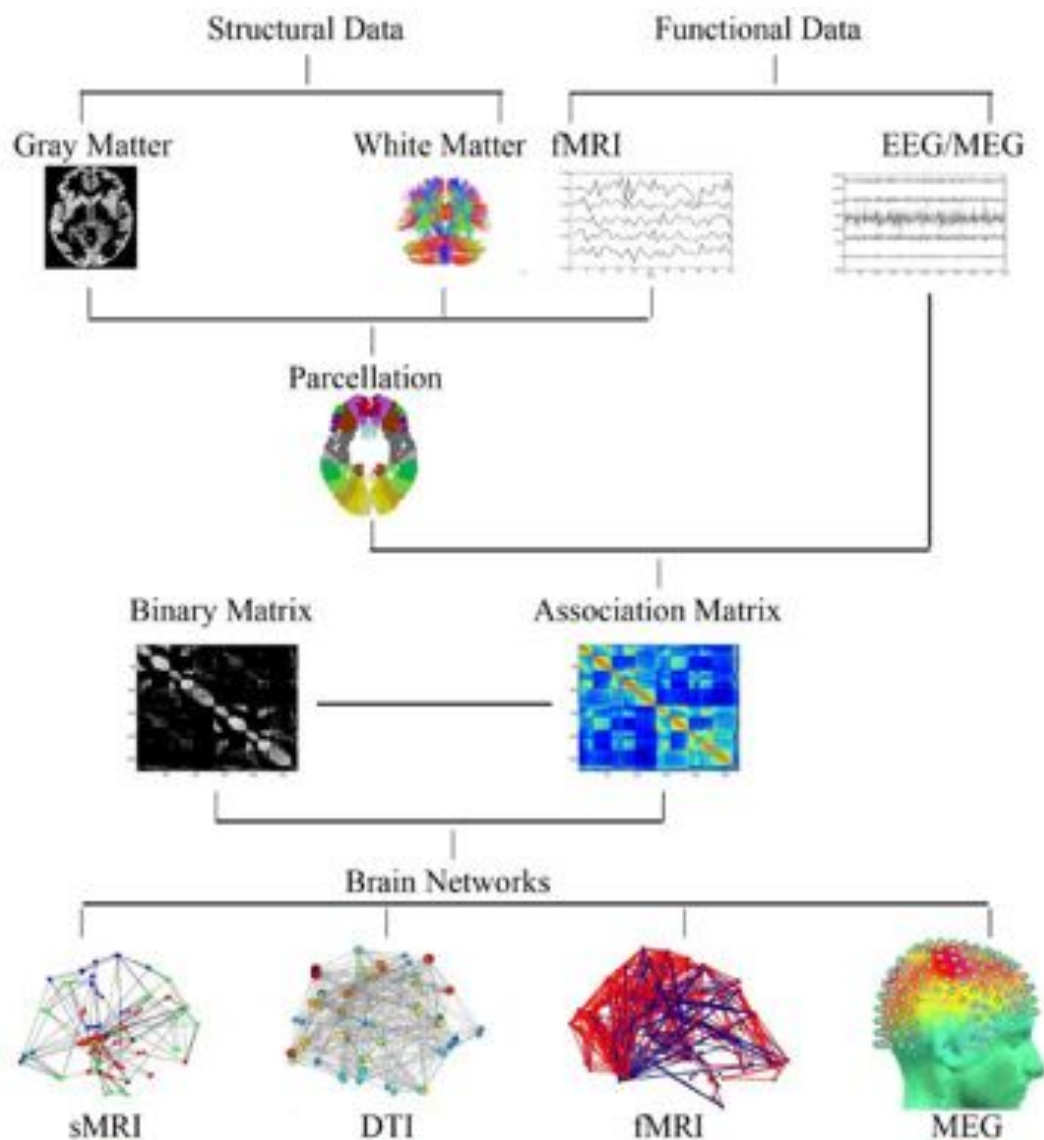


FIG. 1.1 – Aperçu des principales techniques utilisées pour construire le connectome cérébral humain. fMRI = Imagerie par Résonance Magnétique Fonctionnelle, DTI = Imagerie par Tenseur de Diffusion, sMRI = Imagerie par Résonance Magnétique Structurale (anatomique), MEG = Magnéto-EncéphaloGraphie, EEG = Electro-Encéphalographie. (source : Bassett et Bullmore (2009)).

Le support du connectome Le cerveau ou prosencéphale se divise en deux parties : le **télencéphale** constitué des deux hémisphères cérébraux reliés par le corps calleux ainsi que les ganglions de la base et le **diencephale** au centre du télencéphale constitué du thalamus, de l'hypothalamus, des noyaux sous-thalamiques, de l'épiphyse et de l'hypophyse (cf. Fig. 1.2).

De façon schématique, on y distingue trois milieux différents :

- les **ventricules cérébraux**, cavités participant à la sécrétion et à la circulation du **liquide céphalo-rachidien (LCR)** dans lequel baigne l'encéphale. Le LCR a un rôle de protection (amortissement des chocs) et assure la circulation des hormones ;
- la **substance grise** constituée des corps cellulaires des neurones, et constituant le cortex cérébral (composé de deux hémisphères cérébraux reliés par le corps calleux) et les noyaux gris centraux ;
- la **substance blanche** se situant sous le cortex et constituée des prolongements des neurones ou axones, organisés en faisceaux de fibres, qui transfèrent l'information.

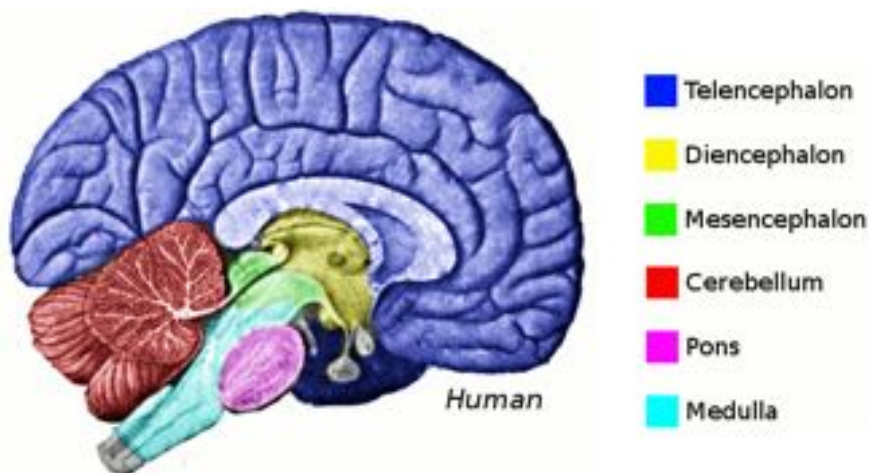


FIG. 1.2 – Les différentes parties de l'encéphale : le cerveau regroupe le télencéphale (bleu) et le diencephale (jaune)

Le cortex ainsi que les noyaux gris, supports de l'activité cérébrale, sont appropriés pour définir le **support du connectome** : les **noeuds du connectome** seront donc définis à l'intérieur des **régions du cortex** ainsi que des **noyaux gris centraux**.

Différentes échelles spatiales peuvent être envisagées pour définir les noeuds du connectome (Sporns et al. 2005, Bohland et al. 2009b). La figure 1.4 en présente un tableau récapitulatif. On reviendra en détails sur les propriétés du connectome dans ces différentes échelles dans la partie 1.2. Nous allons ici décrire brièvement leurs caractéristiques.

Une première **échelle microscopique** peut être définie avec comme noeuds les **corps cellulaires des neurones** et comme connexions leurs terminaisons axonales ou dendritiques communiquant via les **synapses**.

Un tel connectome dont l'élément de base serait le neurone aurait alors un nombre de noeuds estimé à 10^{11} avec 10^{15} connexions. Ces **énormes**

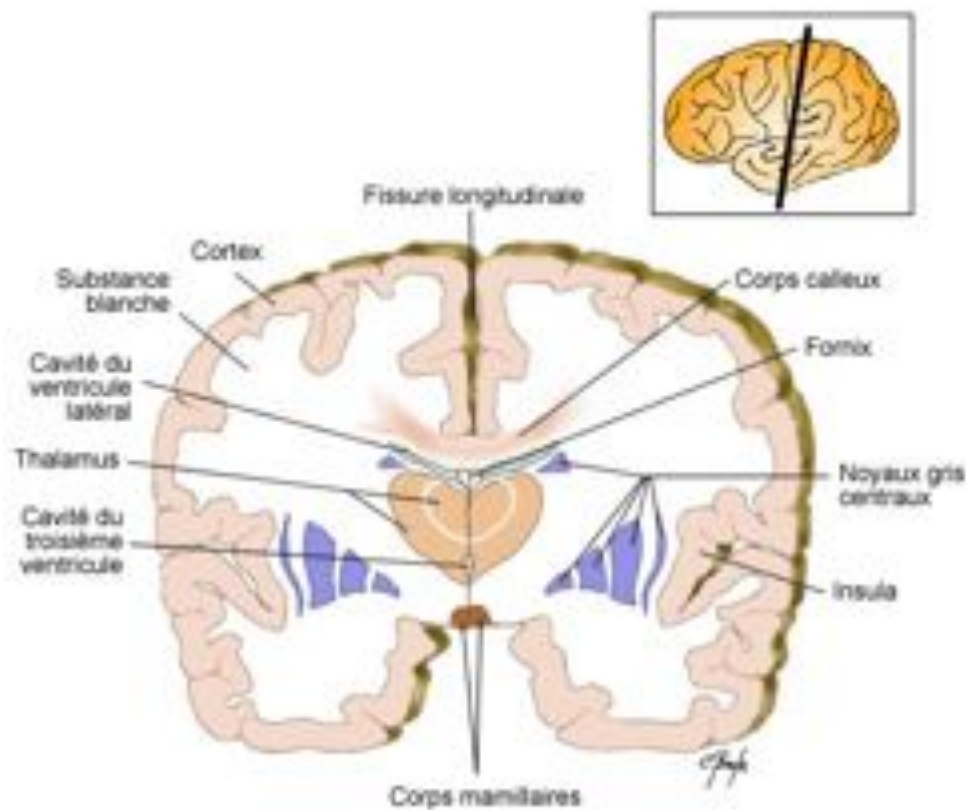


FIG. 1.3 – Les trois milieux constituant le cerveau : ventricules, substance grise, et substance blanche. (source : Chantal Proulx).

dimensions rendent très difficile la construction et l'analyse d'un connectome à cette échelle. L'étude de tels connectomes au niveau d'un groupe d'individus semble en outre compromise car une correspondance à travers les sujets paraît difficile à construire à une échelle si fine.

De plus, contrairement au génome où la mutation génomique d'une paire de base a des conséquences non négligeables, le réseau des connexions cérébrales est beaucoup plus robuste à l'échelle microscopique : le dysfonctionnement d'une unique synapse ou d'un neurone individuel a des conséquences très limitées. Ceci étant notamment dû au fait qu'un processus cognitif fait intervenir une large population de neurones distribués parmi le réseau (Mountcastle 1998). Les éléments de ce connectome évoluent rapidement aussi, résultant de la *plasticité cérébrale* qui s'effectue au niveau du poids attribué à chaque synapse par exemple.

De part cette **complexité**, cette **variabilité interindividuelle**, cette **robustesse** et cette **plasticité cérébrale**, étudier le connectome à l'échelle microscopique ne paraît pas approprié pour l'instant. Ce jugement dépendant bien sûr de l'objectif de l'étude. Une étude microscopique de petits groupes de cellules, comme c'est le cas du projet *Blue Brain* peut s'avérer très intéressante. En effet, ce projet initié par Henry Markram en 2002 à l'Ecole Polytechnique fédérale de Lausanne, a permis d'identifier différents types de neurones en étudiant les quelques milliers de neurones d'une *colonne corticale* (Markram 2006).

A l'opposé, on peut définir une **échelle macroscopique**, où les éléments de base i.e. les noeuds seraient définis à partir de critères macro-anatomiques tels que les **sillons** et les **gyri**, auxquels on rajouterait les **structures sous corticales**. Il existe des **méthodes automatiques** qui construisent déjà les parcellisations associées dans la communauté scientifique (cf. partie 1.2.1.2). Cependant, l'échelle est **trop grossière** pour représenter les motifs plus fins intervenant dans le réseau cérébral lors d'un processus cognitif.

L'**échelle mésoscopique** intermédiaire des deux précédentes consiste à prendre comme éléments de base des **populations de neurones** de même type ou partageant les mêmes propriétés d'organisation. On attend alors une **variabilité inter-individuelle intermédiaire** aussi. Cette échelle paraît être un bon compromis entre l'échelle microscopique et macroscopique. Cependant, le premier **problème est d'identifier les noeuds** du réseau dont la définition n'est pas immédiate. Nous verrons par la suite différents types de données utilisées pour aboutir à de telles parcellisations : des données architectoniques 1.2.1.1, fonctionnelles 1.2.2, de diffusion 1.2.3, ou de connectivité anatomique 1.2.3. Des études de ce type de connectome sont présentes dans Hagmann et al. (2007; 2008). L'un des objectifs de cette thèse est de fournir une méthode d'identification de ces noeuds à partir des données de connectivité anatomique.

Une fois de tels connectomes construits, on peut avoir accès à toute une série de propriétés dérivant de la théorie des graphes, qui seront abordées dans la partie 1.1.4. Nous allons tout d'abord parler des applications associées à l'étude du connectome qui justifient sa popularité.

Echelles et entités associées	Ordre de grandeur	Avantages	Inconvénients
Microscopique environ 10^{11} neurones et 10^{15} synapses	μm	- découpage très fin - entités de base relativement bien définies (neurones, avec en plus les cellules gliales par exemple)	- nécessaire ? (le réseau est robuste face à l'altération d'un neurone ou d'une synapse) - complexité \Rightarrow difficultés pour : (1) la construction et l'analyse, (2) la correspondance entre les individus - dynamique : changements dus à la plasticité cérébrale
Mésoscopique groupes de neurones et connexions les reliant	cm	- découpage intermédiaire à 2 échelles : parcellisation fine (mésoscopique) avec des connexions longue distance à l'échelle macroscopique - plus d'invariance au niveau inter-individuel	- problème concernant la définition des entités de base : pas de processus de parcellisation standard dans la littérature
Macroscopique gyri, sillons \sim 80 régions connexions longue distance	5 cm et plus	- organisation relativement bien connue - outils de parcellisation automatique - relative stabilité inter-individuelle	- découpage grossier : \Rightarrow non capture de motifs plus fins du système cérébral

FIG. 1.4 – Les différentes échelles envisagées pour le connectome avec une liste de leurs principaux avantages et inconvénients.

1.1.2 Vision modulaire et vision holistique

Historiquement, deux manières d'appréhender les processus cognitifs se sont confrontées :

1. **Le paradigme modulaire** pour lequel une fonction cognitive résulte de l'action d'une région cérébrale spécifique qui fonctionne indépendamment des autres régions ; par exemple pour la perception des visages avec l'aire fusiforme (Kanwisher et al. 1997) ou pour la perception du corps humain (Downing et al. 2001) ;
2. **Le "Large scale network" paradigme** (holistique) : qui définit une fonction cognitive comme le résultat de l'action conjointe de régions cérébrales distribuées dans de larges réseaux. Historiquement, ce paradigme a été initié par Luria et al. (1966), Wernicke (1970).

Une accumulation de preuves pèse en faveur d'une représentation holistique des fonctions cognitives d'après Bressler et Menon (2010). De manière plus mesurée, on constate que le paradigme modulaire montre des limitations et des erreurs (Fuster 2000). Le paradigme modulaire suggérerait notamment une parcellisation du cortex en aires sensorielles primaires indépendantes avec une aire par modalité (aire visuelle, auditive ou somato-sensorielle), chacune étant constituée de neurones répondant à une modalité spécifique. Cette vision des aires sensorielles n'est cependant pas si simple, des études ayant montré que ces domaines sont influencés par plusieurs modalités (Wallace February 2004). Récemment, de nombreuses études apportent de nouvelles contributions à la thèse holistique. Parmi celles-ci, nous pouvons citer Freeman (1975), Mountcastle (1978), Goldman-Rakic (1988), Mesulam (1990), Sporns et Zwi (2004), Sporns et al. (2004), Fuster (2005). La confirmation de cette thèse rend évident l'intérêt d'étudier le connectome pour mieux comprendre les processus cognitifs. Dans la partie suivante, nous présentons plus précisément quelques exemples d'applications possibles.

1.1.3 Applications

Ce qui suit est fortement inspiré de Bohland et al. (2009b) dont la lecture est conseillée pour plus de précisions.

Scientifiques

De nombreuses applications de la connaissance du connectome sont envisageables. Parmi les applications scientifiques, nous en présentons ici quelques-unes.

1. **Cognition.** Le connectome est le squelette dans lequel les processus cognitifs prennent forme. Sa connaissance permettrait une meilleure compréhension de ces processus. La compréhension de l'architecture du réseau de connectivité cérébral pourrait en effet permettre de mieux comprendre comment les fonctions cognitives émergent, et quelles sont leurs contraintes. Elle pourrait de plus guider le design expérimental des études électrophysiologiques, pour le placement des électrodes par exemple ainsi que pour mieux interpréter les données acquises par électro-stimulation.

2. **Génétique.** Le connectome que l'on pourrait aussi appeler le *phénotype de connectivité cérébrale* pourrait fournir le lien manquant entre le *phénotype comportemental* et le *génotype* d'un individu.
3. **Evolution des espèces.** Il nous permettrait de plus de mieux appréhender l'évolution des espèces en fournissant un nouvel outil de comparaison. Par exemple la connaissance du connectome mésoscopique a permis de remettre en cause des homologues faites entre des structures télencéphaliques des volatiles et des structures sous corticales des mammifères (Jarvis et al. 2005).

Médicales

Depuis Geschwind (1965) et ensuite Wernicke (1970), le lien entre l'altération des connexions neuronales et les maladies cérébrales est avéré. De nombreuses études ont déjà montré des altérations de la connectivité cérébrale tant pour les maladies développementales que psychiatriques ou dégénératives. En effet, il peut y avoir un développement anormal de la connectivité cérébrale dans des maladies développementales telles que l'**autisme** (Belmonte et al. 2004), avec notamment une réduction du corps calleux et de la connectivité fonctionnelle longue distance (fronto-temporale). Dans des maladies psychiatriques comme la **schizophrénie** (Bullmore et al. 1997, Bassett et al. 2008) ou la **maladie d'Alzheimer** (Buckner et al. 2009), des disconnexions (absence d'une connexion qui est présente chez le sujet sain) ou misconnexions (connexion en plus par rapport au sujet sain) peuvent aussi être mises en évidence.

Il paraît donc essentiel de comprendre le connectome et de bien le représenter pour pouvoir mieux quantifier la sévérité de ces altérations dans la connectivité cérébrale et mieux comprendre les mécanismes sous-jacents. Pour chaque maladie, il convient bien sûr de trouver l'échelle spatiale et la combinaison des différentes mesures de connectivité adaptées pour construire le connectome approprié à l'étude de cette maladie.

La connectivité donne à l'étude de n'importe quelle maladie une dimension en plus : au lieu de considérer chaque région comme indépendante conformément au paradigme modulaire, on étudie l'ensemble des relations entre ces régions, i.e. le réseau sous-jacent. Cette information additionnelle peut alors permettre d'améliorer les études cliniques dans de nombreux aspects, si l'on croit en une organisation holistique du cerveau.

1.1.4 Caractéristiques du connectome

1.1.4.1 Propriétés globales

A travers la théorie des graphes, différentes mesures peuvent être calculées pour décrire le connectome (cf. Annexe B ou Rubinov et Sporns (2009) pour plus de détails). Parmi ces mesures, nous pouvons citer le **degré de clustering**, la **modularité** (mesures de la ségrégation fonctionnelle), ainsi que la **longueur de chemin** (path length) et l'**efficacité de communication** (mesures de l'intégration fonctionnelle).

Récemment, plusieurs études à partir de "tract-tracing" (Sporns et Zwi 2004)), de l'IRM de diffusion (Hagmann et al. 2008, Iturria-Medina et al. 2008, Gong et al. 2008, Bullmore et Sporns 2009), de l'IRM fonctionnelle

au repos (Achard et Bullmore 2007) ou de l'épaisseur corticale (He et Chen 2007) convergent pour reconnaître au connectome d'avoir une organisation de type **small-world**. Ce type de réseau découvert par Watts et Strogatz (1998) est connu pour avoir un **fort degré de clustering** ainsi qu'une **faible longueur de chemin moyenne**, suggérant une organisation en communautés de noeuds localement connectés, chaque communauté partageant des connexions longue-distance avec les autres communautés. Avec cette configuration, les noeuds ne sont pas tous voisins, mais le chemin reliant deux noeuds du réseau est court. Ces réseaux combinent alors des **propriétés d'intégration et de ségrégation fonctionnelles**.

Le connectome cérébral humain a d'autres propriétés, notamment :

- économiques : dans le sens qu'il permet une grande efficacité (globale et locale) dans le traitement en parallèle de l'information, à bas coût i.e. en minimisant les connexions (étude d'IRM fonctionnelle de Achard et Bullmore (2007)),
- une structure composée de hubs : présence de hubs structuraux dans le cortex postérieur médial (études en IRM de diffusion de Hagmann et al. (2008), Gong et al. (2008), et fonctionnelle de Buckner et al. (2009))

1.1.4.2 Exemple d'une étude du connectome

Connectome régional et local Hagmann et al. (2008) a analysé le connectome en prenant comme éléments de base deux segmentations du cortex : une parcellisation en gyri (66 régions) et une sur-segmentation de cette parcellisation en 998 régions d'intérêt. La figure 1.5 récapitule les différentes étapes de la méthode de construction des matrices de connectivité correspondantes. Ces matrices ont été créées pour les deux hémisphères cérébraux pour 5 sujets sans tenir compte des connexions avec les noyaux gris centraux. En résumé, à partir de données de diffusion (haute résolution de $2 \times 2 \times 3 \text{ mm}$) et de données anatomiques (1 mm isotropique), une tractographie est effectuée. Les surfaces corticales sont extraites, mises en correspondance à travers les sujets et segmentées suivant une parcellisation en gyri (à l'aide de Freesurfer, cf. partie 2.3.2.2 pour plus de détails). Chaque région est ensuite sur-segmentée, le tout formant pour un sujet un ensemble de 998 régions compactes d'aire $1,5 \text{ cm}^2$ en moyenne. Les matrices de connectivité entre les régions sont ensuite calculées et moyennées à travers les sujets pour aboutir à une unique matrice moyenne.

Densité de connexion La mesure de connectivité choisie est la densité de connexion qui représente le nombre de tracts par unité de surface de cette connexion. En considérant deux noeuds A et B et leur relation dans le graphe e , ainsi que F_e l'ensemble des tracts reliant A et B , la densité de connexion entre A et B est définie par :

$$C_{\mathcal{A}, \text{density}}^0(A, B) = \frac{2}{\text{surf}_A + \text{surf}_B} \sum_{f \in F_e} \frac{1}{l(f)} \quad (1.1)$$

où n_{tracts} est le nombre de tracts du faisceau, Surf_A et Surf_B étant les aires des parcelles A et B respectivement et $l(f)$ la longueur en mm du tract f .

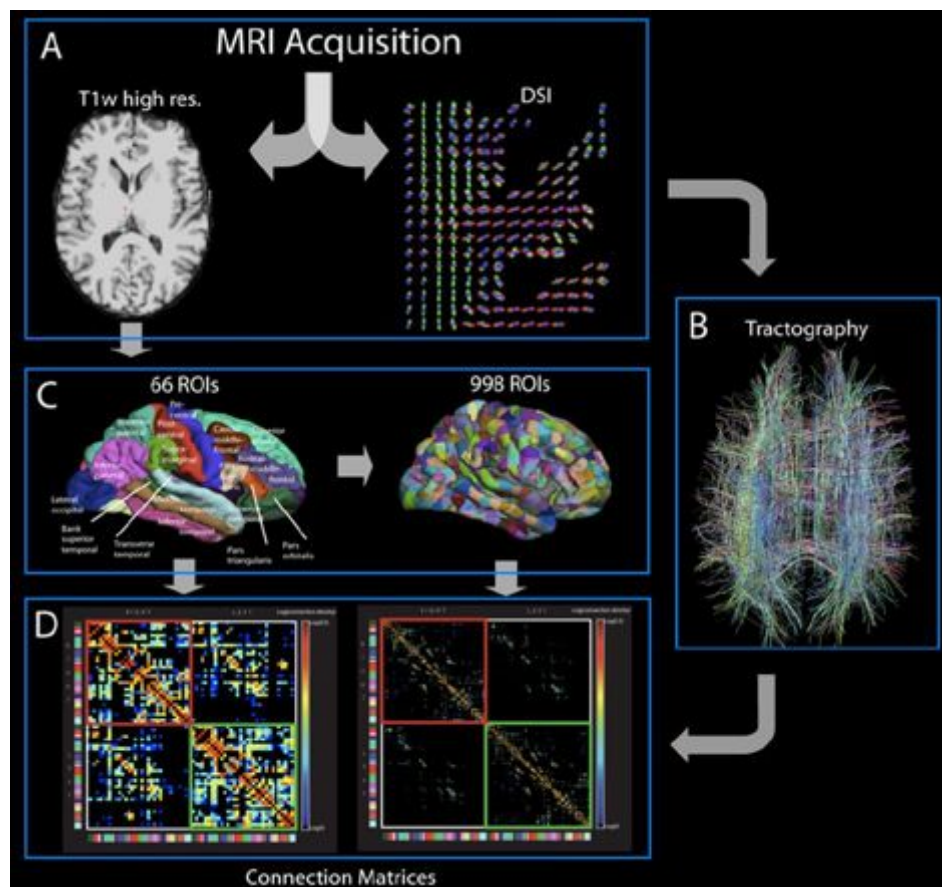


FIG. 1.5 – Construction de matrices de connectivité anatomiques à deux échelles spatiales différentes - les étapes. (A) Acquisition des images IRM de diffusion. (B) Tractographie dans la matière blanche. (C) Segmentation de l'interface Gris/Blanc en régions d'intérêt. (D) Les deux matrices de connectivité créées en utilisant les étapes (B) et (C). (source : Cammoun et al. (2009))

L'unité des poids est alors en nombre de tracts par mm^2 . La matrice obtenue est assez creuse avec des densités de densités de connexion variant entre 2,8 et 3%.

Répartition des connexions Sur la somme totale de toutes les densités de connexion, 54% des connexions relient des régions qui appartiennent à la même région anatomique (au même gyrus), 42% des connexions relient des sous-régions de différentes régions anatomiques (mais dans le même hémisphère), et 4% des connexions sont inter-hémisphériques.

Distribution des degrés et forces de connexion Le degré représente le nombre de connexions par noeud tandis que la force d'un noeud représente la somme des poids de toutes ces connexions. La distribution cumulative des degrés ainsi que celle des forces de connexions trouvée dans cette étude est **exponentielle** (cf. Fig. 1.6 et 1.7). Il y a de plus une variation significative des degrés et forces de connexion à travers les régions avec des régions cumulant degré et force de connexion élevés comme le cortex médial, le cuneus, le precuneus, et le cortex cingulaire antérieur et postérieur. La figure 1.7 montre les forces de connexion des noeuds calcu-

lées à partir des matrices de connectivité haute résolution, pour l'ensemble des sujets. Les régions citées précédemment ont des forces de connexion élevées et de manière reproductible à travers les sujets.

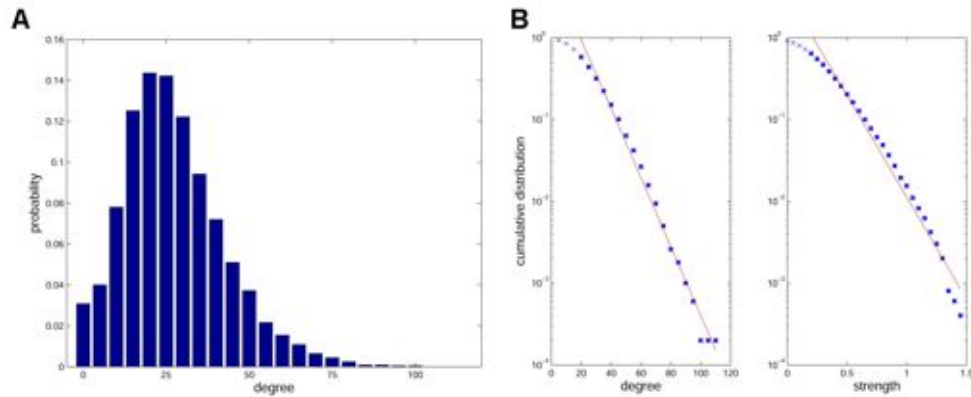


FIG. 1.6 – Distribution des degrés et des forces de connexion sur les matrices de connectivité haute résolution. (A) Distribution des degrés, moyennés à travers les 5 sujets. (B) Graphique log-linéaire des distributions cumulatives des degrés et forces des noeuds agrégés à travers les 5 sujets. La ligne rouge indique la droite de meilleure régression, qui est cohérente avec un déclin exponentiel du degré et de la force de connexion. (source : Hagmann et al. (2008))

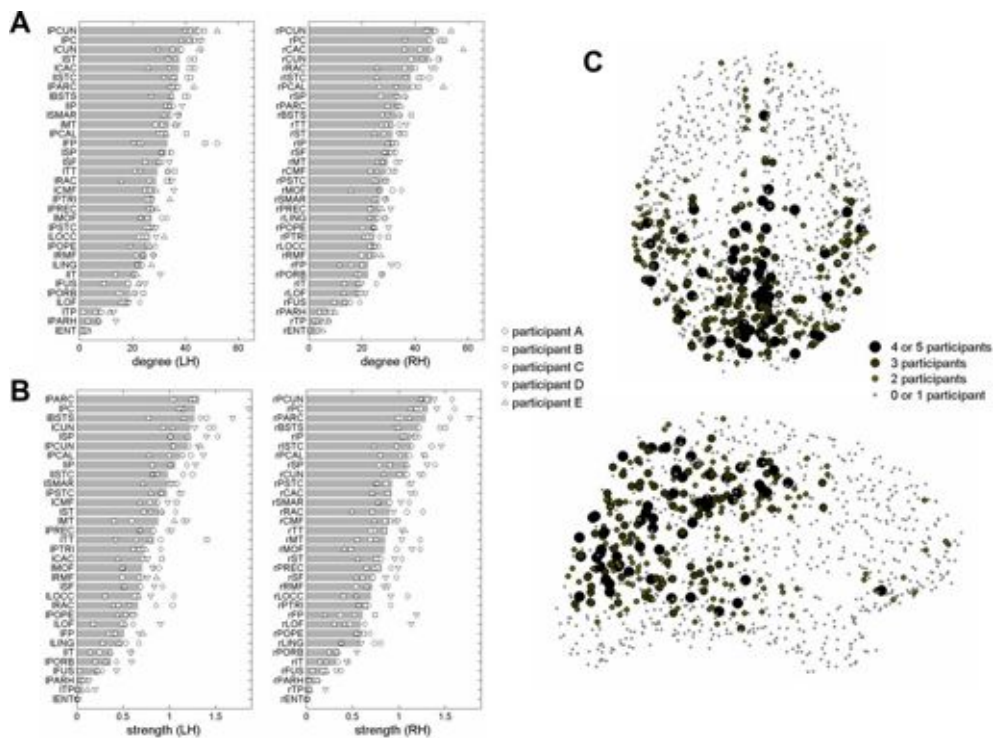


FIG. 1.7 – Distribution des degrés et des forces de connexion. (A) Distribution des degrés de noeuds pour l'hémisphère gauche et droite pour l'ensemble des régions de la matrice de connectivité basse-résolution. Les barres ombrées représentent la moyenne à travers les 5 sujets et les symboles les données individuelles. (B) Idem que (A) mais pour les forces de connexion. (C) Représentation spatiale des forces de connexion des noeuds pour l'ensemble des sujets pour les matrices de connectivité haute résolution. (source : Hagmann et al. (2008))

1.2 PLUSIEURS DÉFINITIONS DES NOEUDS DU CONNECTOME

Etant donné qu'il n'existe pas de méthode de référence pour parcelliser le cortex, les parcellisations utilisées pour étudier le connectome sont nombreuses : à différentes échelles, et avec des nombres de noeuds associés variés (de 100 à 10^4 , cf. Table 1.1). Zalesky et al. (2010) a étudié récemment l'évolution des paramètres globaux précédents en fonction du nombre de noeuds. Il montre par exemple que le caractère small-word ne change pas, mais que son degré d'appartenance à ce type de réseau change.

Etude	Noeuds	Parcellisation	Sujets
Iturria-Medina et al. (2007)	71	IBASM	5
Gong et al. (2008)	78	AAL	80
Iturria-Medina et al. (2007)	90	AAL	20
Hagmann et al. (2007)	500-4 000	Interface Gris/Blanc	2
Hagmann et al. (2008)	998	Interface Gris/Blanc	5

TAB. 1.1 – *Etudes du connectome cérébral humain par IRM de diffusion : Des parcellisations à de multiples échelles. AAL = Automated Anatomical Labeling Atlas, Interface Gris/Blanc = parcellisation aléatoire de l'interface Gris /Blanc (sous forme surfacique), IBASPM = voir Alemán-Gómez et al. (2006). Les informations de ce tableau sont extraites de Zalesky et al. (2010) où le lecteur intéressé trouvera des données concernant des études avec d'autres modalités d'imagerie.*

Parmi les parcellisations utilisées pour définir les noeuds du connectome, nous pouvons citer :

- l'atlas **AAL** (Tzourio-Mazoyer et al. 2002) disponible dans le logiciel **SPM** qui est un atlas de 90 régions dans le repère du MNI, défini à partir de la segmentation d'un sujet en régions d'intérêt à partir de la macro-anatomie du cortex ;
- l'atlas fourni par la boîte à outils **IBASM** (Alemán-Gómez et al. 2006) de 71 régions ;
- Sur-segmentation en 998 sous-régions de tailles similaires d'une segmentation du cortex en 66 régions anatomiques proposée par Free-surfer (Desikan et al. 2006).

Nous allons présenter dans la suite les différentes manières d'obtenir les noeuds du connectome, qui sont associées à des méthodes de parcellisation du cortex. Nous distinguons ces méthodes en trois classes en fonction du type de données qu'elles utilisent :

- les parcellisations à partir de l'anatomie,
- les parcellisations fonctionnelles,
- les parcellisations à partir de la connectivité anatomique.

1.2.1 Les parcellisations à partir de l'anatomie

Les parcellisations structurales se divisent elles-mêmes en deux classes selon l'échelle spatiale de l'analyse (cf. Fig 1.4), qui peut s'effectuer soit au niveau :

1. Microscopique : cyto-architectonie du cortex et autres techniques dont la plupart sont faites ex-vivo,
2. Macroscopique : marqueurs tels que les gyri et les sillons, pour lesquels il existe des procédures de parcellisation automatique in-vivo.

1.2.1.1 Au niveau microscopique

Cytoarchitectonie Le cortex est organisé en couches cytoarchitectoniques dont le nombre et l'épaisseur varient (Fig. 1.8(a)). **Brodmann (1909)** a ainsi proposé une parcellisation du cortex en 43 aires cytoarchitectoniques déterminées à partir du nombre de cellules neuronales, de leur taille et type, ainsi que de leur densité et distribution, mis en évidence par une technique de colorisation des corps cellulaires. Cet atlas, adapté en 3 dimensions par Talairach et Tournoux (Talairach et Tournoux 1988) est très utilisé aujourd'hui dans la localisation corticale.

Economo et Koskinas (1925) et Sarkisov et al. (1949) proposent par la suite d'autres parcellisations suivant les mêmes types de critères, mais qui diffèrent par le nombre, l'étendue, et la configuration des aires trouvées (cf. Fig 1.16).

Toutes ces parcellisations sont faites sur un nombre limité d'individus, et n'étudient pas la variabilité inter-individuelle des aires obtenues.

Des parcellisations **myéloarchitectoniques** ont aussi été effectuées à partir de différences de myélinisation du cortex (Smith 1907, Vogt et Vogt 1919) (cf. Fig 1.16). Plus récemment des méthodes automatiques et indépendantes de l'observateur ont été développées (Annese et al. 2004, Schleicher et al. 1999). Il est aussi possible de cartographier les **récepteurs neuronaux** et de parcelliser le cortex en conséquence (Zilles et Amunts 2009)

Ces méthodes **ex-vivo** sont difficiles à mettre en oeuvre et les parcellisations obtenues couvrent seulement un petit pourcentage du cortex. Leur transposition in-vivo par normalisation sur des atlas probabilistes est une alternative (cf. partie 1.4.2).

Récemment, l'arrivée des IRM à hauts champs ouvre de nouvelles perspectives, avec ainsi des essais d'**histologie in-vivo** (myéline Walters et al. (2007), avec le signal de phase Duyn et al. (2007)).

1.2.1.2 Macro-Anatomie

Au niveau macroscopique la surface du cortex est plissée. Les circonvolutions du cortex sont constituées de vallées (ou plis) : les **sillons**, et de sommets : les **gyri**. A partir de cette anatomie sulco-gyrale, plusieurs types de parcellisations surfaciques en gyri du cortex et individuelles (dans l'espace natif du sujet) ont été développées.

Cachia et al. (2003) propose ainsi une parcellisation en gyri faite à partir de l'identification automatique des sillons du cortex (Rivière et al. 2002), le fond des sillons définissant la frontière entre deux gyri. Clouchoux et al. (2006) délimite les gyri du cortex à partir de l'identification de quelques sillons et d'un système de coordonnées surfaciques associé (Clouchoux et al. 2005). Ces deux méthodes sont dans Brainvisa¹, logiciel de traitement d'images cérébrales. Dans Freesurfer², Desikan et al. (2006) construit d'abord un atlas à partir de 40 sujets dont les gyri ont été étiquetés manuellement et parcellise ensuite de nouveaux sujets à partir de cette base d'apprentissage.

¹<http://brainvisa.info>, cf. 2.3.1 pour plus de détails.

²<http://surfer.nmr.mgh.harvard.edu/>, cf. 2.3.1 pour plus de détails.

Il est important de préciser que dans ces méthodes, dans certaines zones, des **délimitations arbitraires sont parfois effectuées**. La figure 1.8(b) illustre bien cela en montrant la parcellisation sur une sphère : l'organisation en méridiens et parallèles suit les fonds de sillons quand il y a des sillons reconnus, mais ailleurs il n'y a aucune contrainte.

1.2.2 Les parcellisations à partir des données fonctionnelles

Historiquement, la découverte et la définition des aires fonctionnelles s'est faite par rapport à des études ex-vivo regardant le lien entre lésions du cerveau et troubles cognitifs associés (aire de Broca 1861 et de Wernicke (1970), cf. Fig. 1.9).

Grâce à des techniques de plus en plus sophistiquées comme la Tomographie par Emission de Positron (PET) ou l'Imagerie par Résonance Magnétique Fonctionnelle (fMRI), il est maintenant possible d'avoir accès à une mesure indirecte de l'activation cérébrale associée à une tâche cognitive.

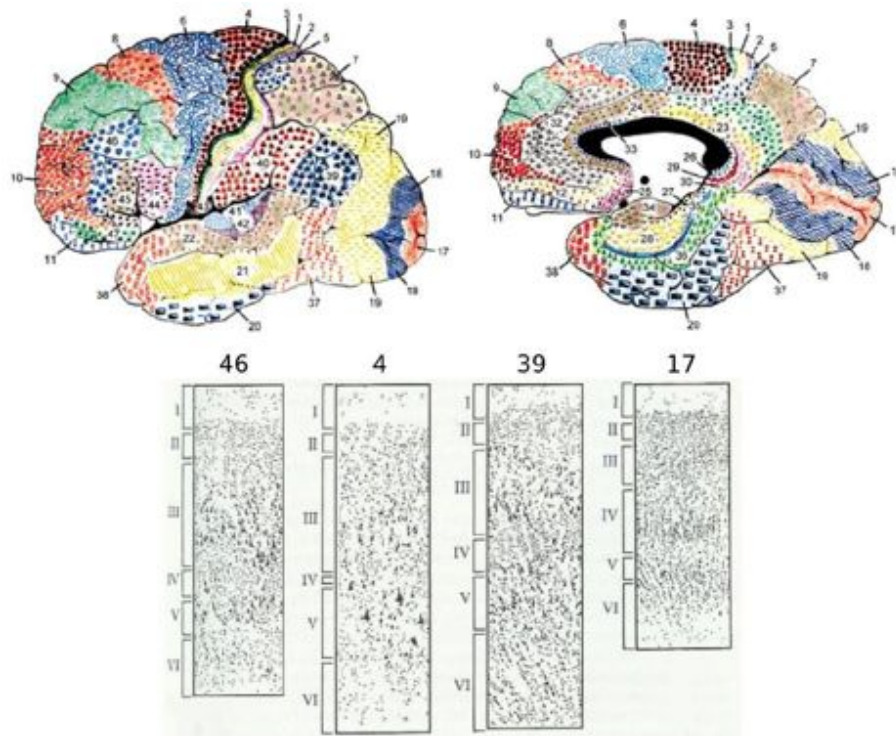
L'IRM fonctionnelle avec tâches définit en chaque élément de volume (voxel) du cortex son "degré" d'activité en réponse à différentes tâches cognitives. Il est alors possible de parcelliser le cortex en fonction du profil d'activité estimé.

Neumann et al. (2006) a proposé ainsi une méthode reposant sur la dynamique de reproduction qui subdivise une région corticale prédéfinie (cortex frontal latéral gauche) de manière reproductible à travers les sujets. Thirion et al. (2006) utilise un algorithme de clustering spectral et parcellise le cortex en régions qui ont des positions proches ainsi que des activités fonctionnelles similaires. Golland et al. (2007) présente également une parcellisation sans a priori grossière, mais consistante.

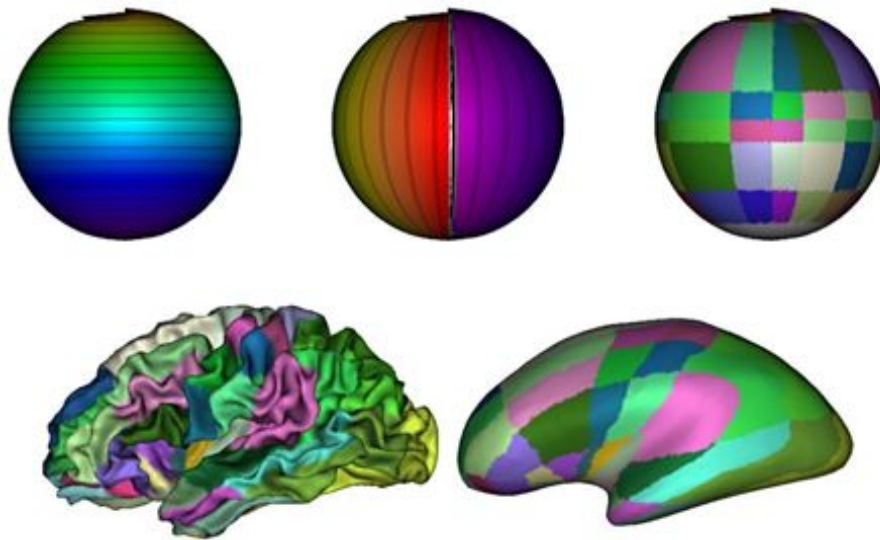
Des études surfaciques liant imagerie fonctionnelle et structurelle ont aussi conduit à des parcellisations (Flandin et al. 2002; 2003, Tucholka et al. 2008).

L'IRM fonctionnelle au repos Il est aussi possible de définir des régions du cortex en fonction de la connectivité fonctionnelle extraite de données d'IRM fonctionnelle au repos ("resting-state functional connectivity"). La connectivité fonctionnelle permet de trouver les régions qui coactivent entre elles, en l'absence de tâche. Cette information peut ainsi être utilisée pour déterminer des réseaux de connectivité. Dans la littérature, Gusnard et Raichle (2001), Raichle et al. (2001), Raichle et Snyder (2007) ont ainsi mis en évidence le **Default-mode network**.

Ces données de connectivité peuvent aussi être clusterisées pour donner des régions fonctionnelles (Cohen et al. 2008, Karkar et al. 2009).



(a) Aires de Brodmann et organisation en 6 couches du cortex : I. Moléculaire, II. Granulaire externe, III. Pyramidale externe, IV. Granulaire interne, V. Pyramidale interne, et VI Polymorphe.



(b) Parcellisation en gyri de Clouchoux et al. (2006) effectuée à partir d'un système de coordonnées surfaciques et de l'identification automatique de quelques sillons.

FIG. 1.8 – Echelles micro et macro-scopique, deux niveaux d'organisation. 1.8(a). Microscopique en aires cytoarchitectoniques, 1.8(b). Macroscopique en circonvolutions composées de gyri et sillons identifiés ici automatiquement. (Images modifiées d'après Tucholka (2010)).

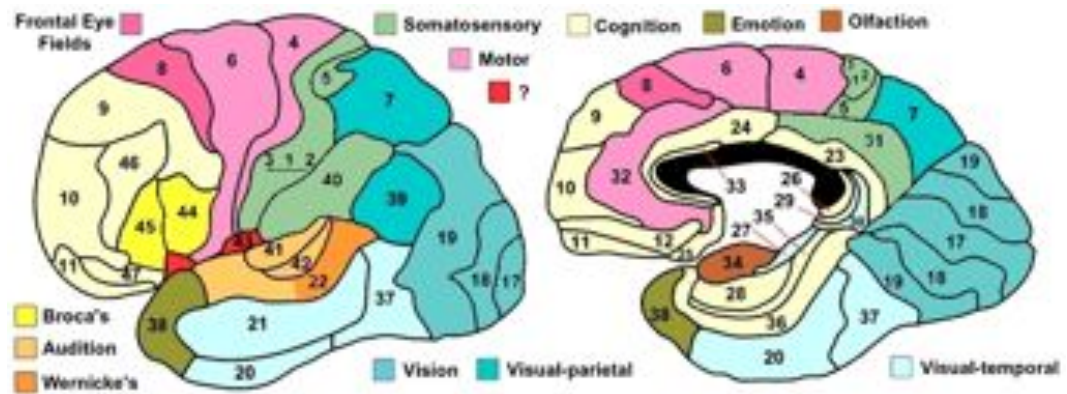


FIG. 1.9 – Aires de Brodmann avec leur caractérisation fonctionnelle. Figure modifiée d'après l'originale de Mark Dubin.

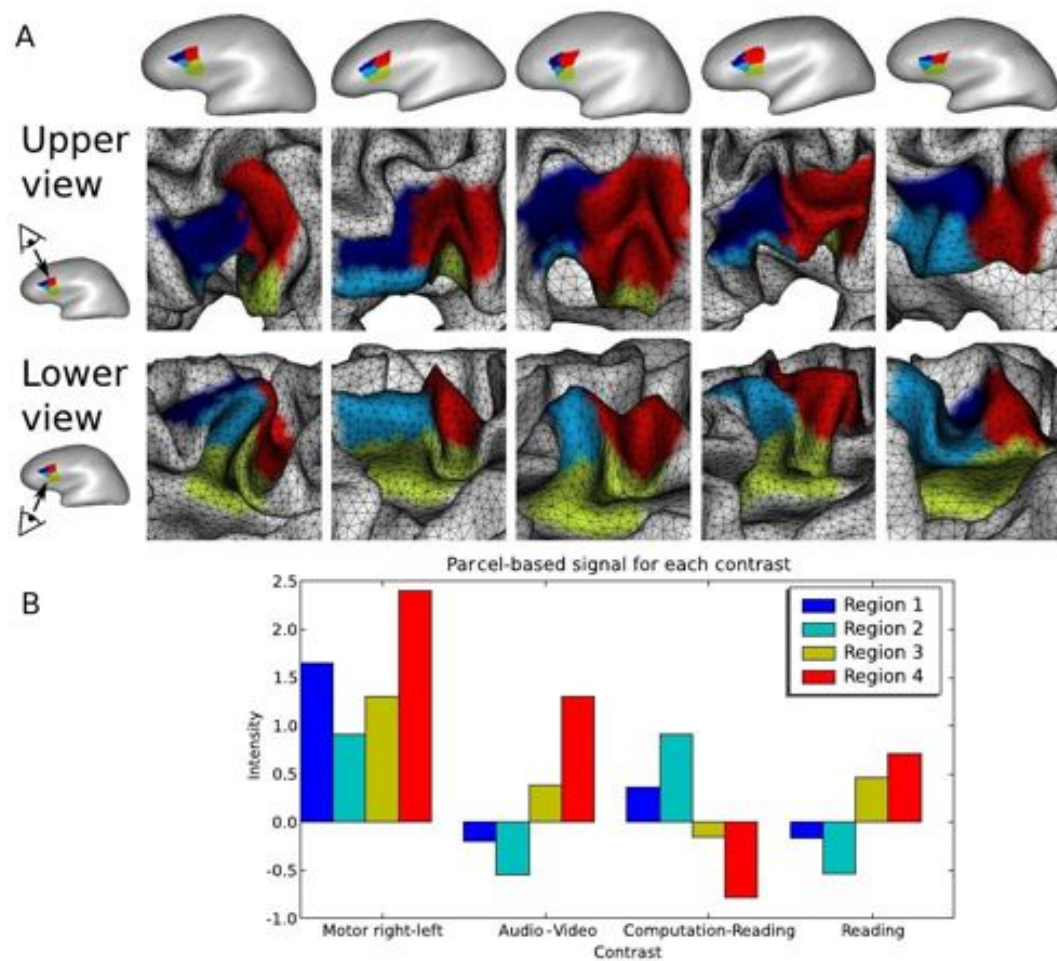


FIG. 1.10 – Parcellisation de l'aire de Broca à partir de données d'IRM fonctionnelle (A) et activités associées à chaque région trouvée (B). Image modifiée d'après Tucholka et al. (2008).

1.2.3 Les parcellisations à partir de la connectivité anatomique

1.2.3.1 IRM de diffusion et hypothèses

L'**IRM de diffusion** est une modalité d'imagerie sensible au mouvement des molécules d'eau dans le cerveau qui en est composé à 80%. Cette modalité permet aujourd'hui de **reconstruire de manière non-invasive l'architecture de la substance blanche** et le réseau de connectivité sous-jacent. L'organisation de la matière blanche en faisceaux de fibres induit en effet une anisotropie du mouvement local des molécules d'eau. L'IRM de diffusion permet alors de collecter cette information.

A partir de ces données, il est possible de tractographier les faisceaux de fibres. Le principe de la tractographie étant de partir d'un voxel "graine" et de reconstruire pas à pas le tract³ passant par ce voxel en utilisant les informations données par l'IRM de diffusion. Un masque de tracking indiquant les voxels cibles pouvant être atteints par un tract est défini préalablement (cf. parties 2.2.4 et 2.3.3 pour plus de détails sur les principes de l'IRM de diffusion ainsi que les algorithmes de tractographie).

Combinée à l'IRM anatomique qui permet de reconstruire le cortex et de séparer matière grise et matière blanche, l'IRM de diffusion permet ainsi d'accéder à la connectivité anatomique du cortex cérébral (cf. parties 2.3.2 et 3.3 décrivant l'extraction du cortex ainsi que la construction de la matrice de connectivité associée).

Suivant l'hypothèse qu'une **région fonctionnelle cérébrale** a un **pattern de connectivité spécifique** (Passingham et al. 2002), la connectivité anatomique cérébrale apparaît alors comme le candidat idéal pour déterminer les noeuds d'un connectome dont les connexions seraient déterminées par cette même connectivité. L'hypothèse sous-jacente étant que la fonction d'une région est définie par ses voies de communication d'entrée et de sortie, à savoir ses connexions.

1.2.3.2 Exemple de parcellisations

Behrens et al. (2003) parcellise ainsi le thalamus en fonction de sa connectivité vers des régions du cortex. De nombreuses études s'ensuivent, parcellisant différentes régions du cortex comme le cortex moteur latéral (Tomassini et al. 2007, Schubotz et al. 2010), l'aire motrice supplémentaire SMA versus pré-SMA (Johansen-Berg et al. 2004, Klein et al. 2007, Anwander et al. 2007, Nanetti et al. 2009, Jbabdi et al. 2009, Crippa et al. 2010), la région de Broca (Anwander et al. 2007, Klein et al. 2007), l'insula (Nanetti et al. 2009), le cortex cingulaire (Beckmann et al. 2009), le gyrus post-central (Guevara et al. 2008), et même le cortex entier (Perrin et al. 2008). Le tableau A de l'annexe A) récapitule pour chacune de ces études leurs principales caractéristiques.

La plupart des approches sont **volumiques** et utilisent un tracking probabiliste pour reconstruire les tracts. Les voxels "graines" sont définis et correspondent à la région à parcelliser.

Dans l'**approche aveugle**, Les **voxels "cibles"** sont **tous les voxels de la matière blanche**. Une tractographie probabiliste initiée à partir des

³Nous appelons tract les fibres reconstruites par tractographie, qui sont des approximations numériques des fibres réelles.

voxels graines permet alors d'avoir pour chaque voxel graine sa probabilité de connexion vers tous les voxels cibles. L'ensemble de ces "forces de connexion" est appelé **tractogramme** (Anwander et al. 2007). Ces probabilités sont soit binarisées soit subissent d'autres opérations (pour prendre en compte la longueur des tracts notamment). La matrice de corrélation entre ces tractogrammes est ensuite calculée, et un clustering (k-means dans la plupart des études) est effectué à partir de cette matrice, avec un nombre de clusters fixé.

Dans l'**approche par régions d'intérêt**, les régions cibles ne sont plus des voxels mais des régions définies manuellement d'après des a priori tirés d'études de "tracing" chez le primate (Rushworth et al. 2005, Beckmann et al. 2009) ou à partir d'une segmentation a priori à partir de critères macro-anatomiques ou fonctionnels (Perrin et al. 2008, Guevara et al. 2008, Jbabdi et al. 2009). On peut ensuite clusteriser la région à partir des profils de connectivité directement.

Exemple du SMA/pré-SMA

Dans cette partie, nous récapitulons les résultats trouvés dans la littérature à propos de la parcellisation de la région du cortex composée du SMA (Aire motrice Supplémentaire) et du pré-SMA.

Rappelons que chez le macaque, la partie du cortex frontal médian correspondant à l'aire 6 de Brodmann chez l'homme peut se diviser en deux aires cytoarchitectoniques : le SMA et le pré-SMA (Vogt et Vogt 1919, Mattelli et al. 1991). Ces deux régions diffèrent aussi par leurs connexions ainsi que par leur fonction (Goldberg 1985, Tanji 2001), le SMA jouant un rôle dans l'organisation des mouvements et présentant une représentation somatotopique du corps (comme le cortex moteur). Le pré-SMA aurait un rôle plus abstrait (Picard et Strick 1996) dans la planification des mouvements ainsi que dans l'acquisition de nouvelles compétences motrices.

Chez l'homme, des études cytoarchitectoniques ont ainsi divisé l'aire 6 de Brodmann en 2 (Zilles et al. 1996) ou 3 régions (Vorobiev et al. 1998).

Aujourd'hui, des parcellisations de cette zone à partir de la connectivité anatomique ont été effectuées (Johansen-Berg et al. 2004, Klein et al. 2007, Anwander et al. 2007, Nanetti et al. 2009), aboutissant à **deux ou trois régions** également. Dans ces approches (cf. Table A de l'annexe A), à partir d'une tractographie probabiliste, une mesure de **connectivité** entre tous les voxels de la région à parcelliser (réunissant le SMA et le pré-SMA, cf. Fig. 1.11) et **l'ensemble des voxels du cerveau** (ou de la matière blanche) est calculée. Chaque voxel de la région à parcelliser a ainsi un profil de connectivité qui représente l'ensemble de ces connexions (probabilité binarisée ou non), ce profil est appelé **tractogramme**. A partir de la matrice de corrélation entre ces profils de connectivité, il est alors possible de clusteriser cette région (utilisation principalement du **Spectral reordering** ou du **k-means**).

Jbabdi et al. (2009) utilisent une approche plus sophistiquée (clustering à partir d'un modèle de mélange hiérarchique infini) qui permet de déterminer le nombre de clusters ainsi que d'effectuer le clustering sur plusieurs sujets simultanément. Il faut noter qu'alors les **profils de connec-**

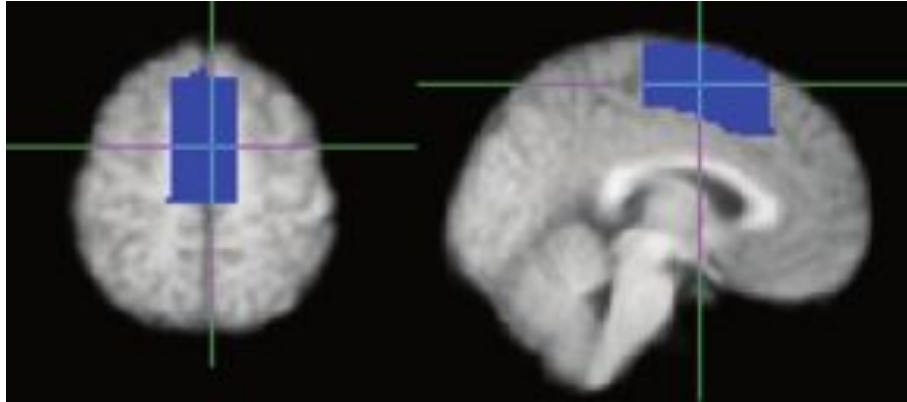


FIG. 1.11 – Définition de la région englobant le SMA et le pré-SMA. Masque utilisé pour définir la région d'intérêt en vue axiale (gauche, $Z = 58$) et sagittale (droite, $X = 2$). La ligne verticale indique $Y = 0$ (ligne verticale à partir de la commissure antérieure, appelée VAC line en anglais). Les coordonnées sont indiquées dans le repère de Talairach. (source : Johansen-Berg et al. (2004)).

tivité utilisés sont réduits vers un ensemble de régions d'intérêt définies par une segmentation a priori du cortex (en 7 ou 48 régions). Il trouve deux régions principales.

Récemment l'étude de Crippa et al. (2010) identifie une **transition graduelle** entre le SMA et le pré-SMA en utilisant un algorithme de dessin de graphe à partir de forces (force-directed graph layout en anglais).

Les figures 1.13, 1.14 et 1.12 montrent les résultats obtenus respectivement par Johansen-Berg et al. (2004) et Anwander et al. (2007). Ces deux études trouvent des résultats cohérents. La première trouve 2 régions, la deuxième 3 régions, l'une étant attribuée au cortex moteur.

Les deux régions correspondant au **SMA et pré-SMA** présentent des **connexions distinctes**. Le SMA est principalement connecté au gyrus pré-central et au tract corticospinal, tandis que le pré-SMA est connecté au cortex préfrontal (avec une terminaison dans le gyrus frontal supérieur) ainsi qu'au cortex pariétal médial.

1.2.3.3 Les deux types d'approches abordées

Dans ces travaux, deux approches pour traiter la connectivité anatomique sont abordées, en fonction du profil de connectivité utilisé pour regrouper les points du cortex ayant des connexions similaires :

1. l'approche **aveugle**, où le **profil de connectivité vers l'ensemble du cerveau** ou de la matière blanche est considéré pour la parcellisation (cf. Table A de l'annexe A) ;
2. l'**approche par régions d'intérêt** : où les **connexions** sont calculées **vers un ensemble de régions d'intérêt**.

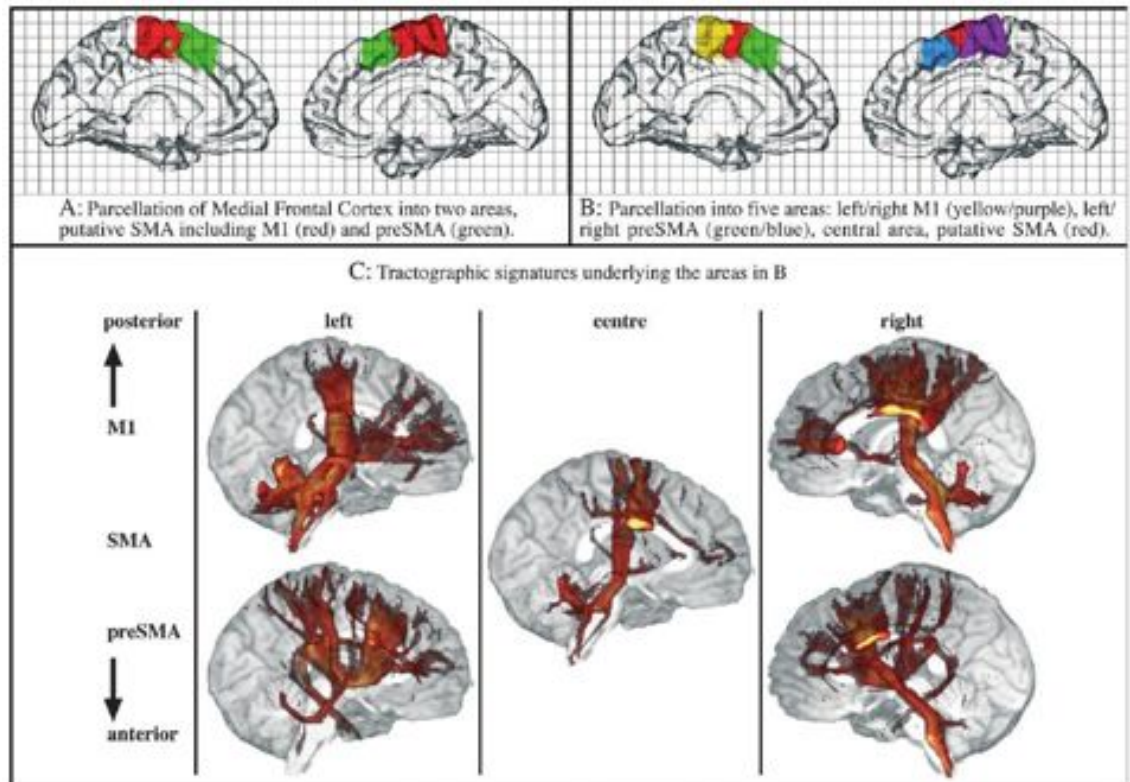


FIG. 1.12 – Parcellisation du cortex frontal médial pour un sujet. (A) et (B) : La parcellisation volumique est projetée sur les surfaces du cortex médial pour les hémisphères droits et gauches. (C) Connexions des aires trouvées (B). (source : Anwender et al. (2007)).

1.3 PLUSIEURS DÉFINITIONS DES CONNEXIONS DU CONNEXOTOME

Nous avons vu précédemment différentes possibilités pour définir les noeuds du connectome. Associées à celles-ci, il y a également différentes façons de construire les connexions du connectome.

Nous pouvons ainsi citer la **connectivité fonctionnelle** (Perlberg et Marrelec 2008), qui comme nous l'avons vu précédemment, consiste à trouver les régions qui co-activent lors d'une expérience d'**IRM fonctionnelle** au repos (sans tâche).

Les **corrélations entre des mesures structurelles** (He et Chen 2007), comme l'épaisseur du cortex ou la courbure, peuvent aussi servir de mesure de "connectivité" entre des zones du cortex. La modalité d'IRM correspondante est l'**anatomie**.

La **connectivité anatomique** (Hagmann et al. 2008) est une mesure de plus. Cette connectivité structurelle, reflet de l'architecture de la substance blanche peut être reconstruite in vivo à partir de l'**IRM de diffusion** suivi d'un algorithme de tractographie. Nous verrons à la partie 2 plus en détails cette reconstruction.

Dans cette thèse nous nous intéresserons à l'étude d'un connectome dont les noeuds et les connexions seront à définir en fonction de la connectivité anatomique.

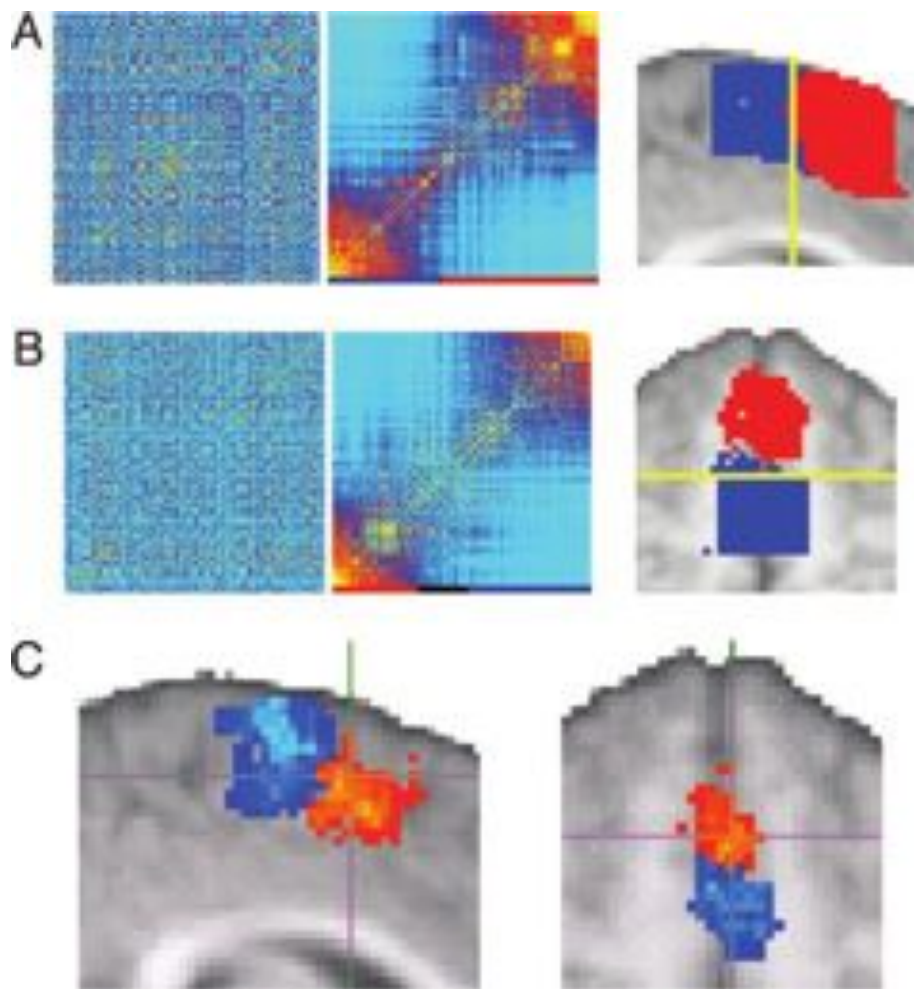


FIG. 1.13 – Parcélisation du cortex frontal médial à partir de la connectivité. Résultats pour la parcellisation (A) d'une coupe sagittale ou (B) axiale, avec la matrice de corrélation des profils de connectivité originale (à gauche) et réordonnée (à droite). (C) Cartes probabilistes des régions. (source : Johansen-Berg et al. (2004)).

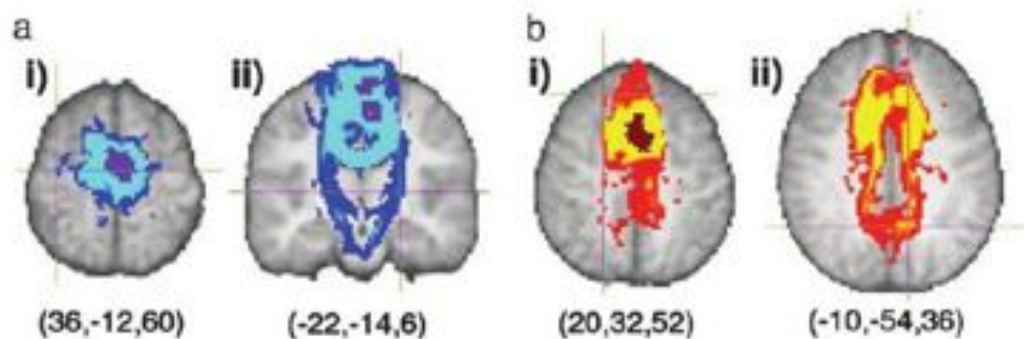


FIG. 1.14 – Connexions du SMA (a) en bleu et du pré-SMA (b) en rouge. Les connexions du SMA vont en direction du cortex moteur (gyrus précentral) et le faisceau cortico-spinal. Les connexions du pré-SMA sont dirigées vers le cortex frontal ainsi que le cortex pariétal médial. (Figure modifiée d'après Johansen-Berg et al. (2004)).

1.4 CONNECTOME ET VARIABILITÉ INTER-INDIVIDUELLE

La variabilité du connectome peut se décomposer en différents types de variabilité inter-dépendants :

- la variabilité du support du connectome, i.e. le cortex,
- la variabilité des aires fonctionnelles (cytoarchitectoniques ou autres),
- la variabilité de l'activité cérébrale,
- la variabilité des connexions sous-jacentes.

Un exemple d'inter-dépendance est la relation entre la variabilité des aires corticales et celle des connexions les reliant : la variabilité dans la position et la trajectoire des faisceaux reliant deux aires fonctionnelles est intimement liée à la variabilité de position de ces aires.

1.4.1 Variabilité du cortex : Support des noeuds du connectome

La surface de chaque hémisphère couvre en moyenne 900 cm^2 . Elle est modelée de façon complexe, formant des plis dont les vallées sont appelées **sillons** et les sommets des **gyri**. Cette anatomie sulco-gyrale est extrêmement variable d'un individu à un autre. On peut distinguer différents types de sillons, en fonction de leur ordre d'apparition à la surface du cortex lors du développement du fœtus et de l'enfant, c'est à dire lors de la **sulcogénèse** : les primaires apparaissant en premier, tels que le sillon central, les secondaires, tertiaires, etc. Plus l'âge d'apparition d'un sillon est élevé, plus sa variabilité interindividuelle est grande. Ainsi, les sillons primaires sont les plus stables alors que les sillons tertiaires sont très variables. La figure 1.15 illustre bien cette variabilité. Elle représente 3 cerveaux où les sillons ont été étiquetés manuellement, avec l'attribution d'une couleur par sillon.

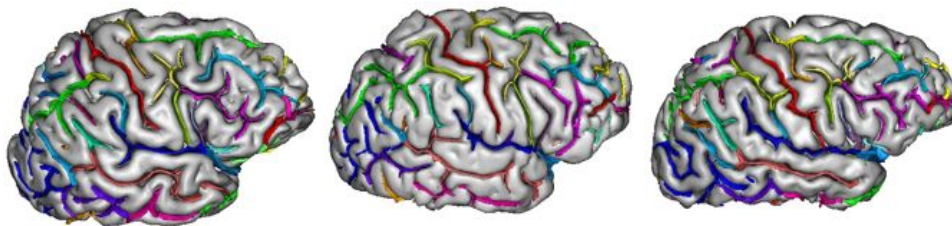
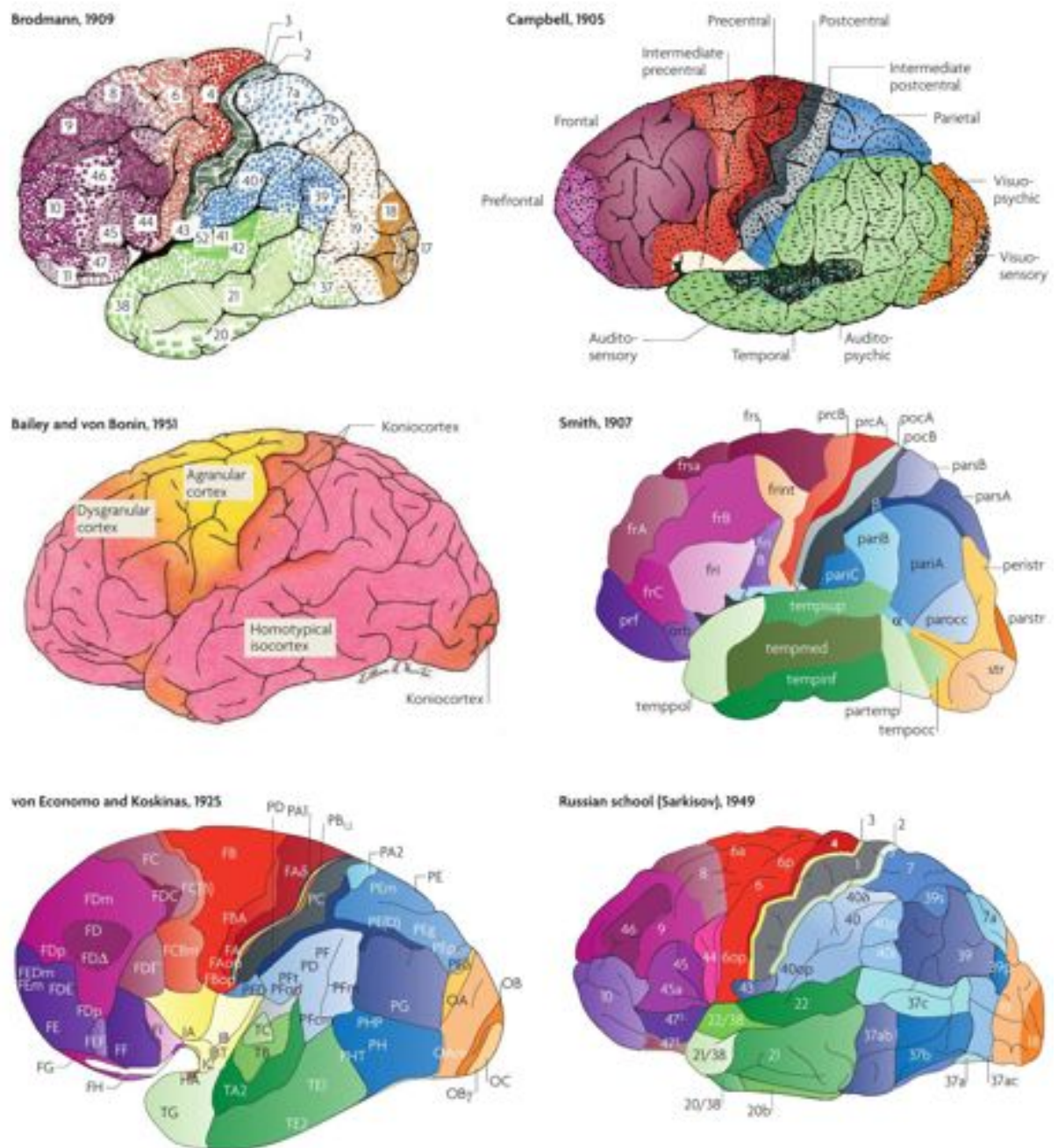


FIG. 1.15 – Cortex de trois individus avec une labélisation manuelle des sillons. Une couleur = un sillon. Exemple de sillon primaire : le sillon central en rouge ; Exemples de sillons apparaissant plus tardivement : sillon bleu et sillon violet du cortex frontal.

Les causes de cette variabilité sont encore mal expliquées, ainsi que les phénomènes liés à la sulcogénèse. Une des hypothèses est que la formation des sillons résulterait de la tension exercée par les faisceaux de fibres sur le cortex pendant leur développement (Van Essen 1997).

1.4.2 Variabilité des modules "fonctionnels" : Les noeuds

Variabilité de l'activité cérébrale A la variabilité anatomique s'ajoute une variabilité fonctionnelle. Nous allons seulement en présenter un aperçu rapide ici, le lecteur intéressé pourra se reporter à la thèse



d’Operto (2008) ou à Thirion et al. (2007) pour plus de détails, notamment sur l’analyse de groupe en IRM fonctionnelle.

Les données d'IRM fonctionnelle présentent une très grande variabilité interindividuelle et une variabilité intra-sujet (ou encore inter-session) plus faible (Smith et al. 2005). Les causes de cette variabilité sont complexes et multiples. Comme facteurs liés à la variabilité fonctionnelle, nous pouvons citer :

- la variabilité anatomique des structures corticales sous-jacentes, le côté macroscopique ayant été décrit précédemment ;

- la variabilité spatiale des “aires fonctionnelles”, illustrée par la suite dans 1.4.2
- les facteurs émotionnels comme le stress ou la fatigue,
- les facteurs cognitifs : utilisation de différentes stratégies pour réaliser une tâche,
- la réponse hémodynamique dont la forme varie spatialement, alors que les analyses classiques en IRMf n’en tiennent pas compte (Ciuciu et al. 2003).

La variabilité observée s’exprime de différentes manières, avec par exemple :

- l’occurrence ou l’absence de certaines activations,
- une variation en intensité et étendue spatiale de l’activation,
- une différence de localisation de cette activation avec un décalage spatial sur la surface corticale.

Variabilité des parcellisations du cortex Cette variabilité spatiale peut être due à la variabilité spatiale des “aires fonctionnelles” sous-jacentes dont la définition ainsi que l’identification sont multiples comme nous l’avons vu précédemment à la partie 1.2.

Ainsi, de nombreuses parcellisations définies à partir de la cytoarchitectonie du cortex ont été proposées. La figure 1.16 en montre quelques-unes. Ces parcellisations sont très différentes. Cette variabilité est en partie causée par la subjectivité des techniques de parcellisation utilisées qui sont dépendantes des observateurs, ainsi que par une différence dans les concepts de l’organisation corticale. Elle est aussi le reflet de la variabilité inter-individuelle, chaque parcellisation ayant été obtenue à partir d’individus distincts.

Les atlas probabilistes

Plus récemment des cartes probabilistes de la cytoarchitectonie ont été créées (Amunts et Zilles 2001, Eickhoff et al. 2005; 2007, Schleicher et al. 2009). Elles permettent de quantifier la variabilité inter-individuelle de ces régions. Ces atlas volumiques couvrent une partie non complète du cortex (cf. Fig. 1.17). Un recalage est nécessaire pour transformer les cerveaux de chaque individu vers un cerveau “template” pour construire cet atlas. Dans ce contexte, les techniques volumiques ne tiennent pas compte de l’anatomie sulco-gyrale sous-jacente. L’approche surfacique, elle, tient compte des sillons et des gyri. Des cartes probabilistes surfaciques sont proposées par Freesurfer pour quelques aires cytoarchitectoniques (Fischl et al. 2007) (cf. Fig. 1.18).

1.4.3 Variabilité des connexions : cas anatomique

A ma connaissance, peu de travaux quantifiant la variabilité des connexions anatomiques ont été effectués que ce soit par des techniques post-mortem ou in-vivo.

En post-mortem, l’identification microscopique de tracts et le recalage 3D de ces données pour un ensemble de sujets donne de précieuses informations sur la variabilité inter-sujets et un moyen d’évaluation des tracts

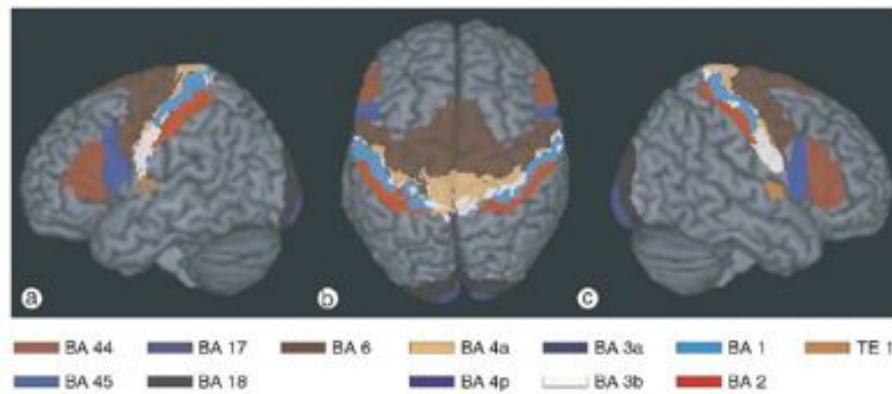


FIG. 1.17 – Les différentes régions couvertes par l'atlas proposé par Eickhoff et al. (2005) dans SPM. Image tirée de cette publication.

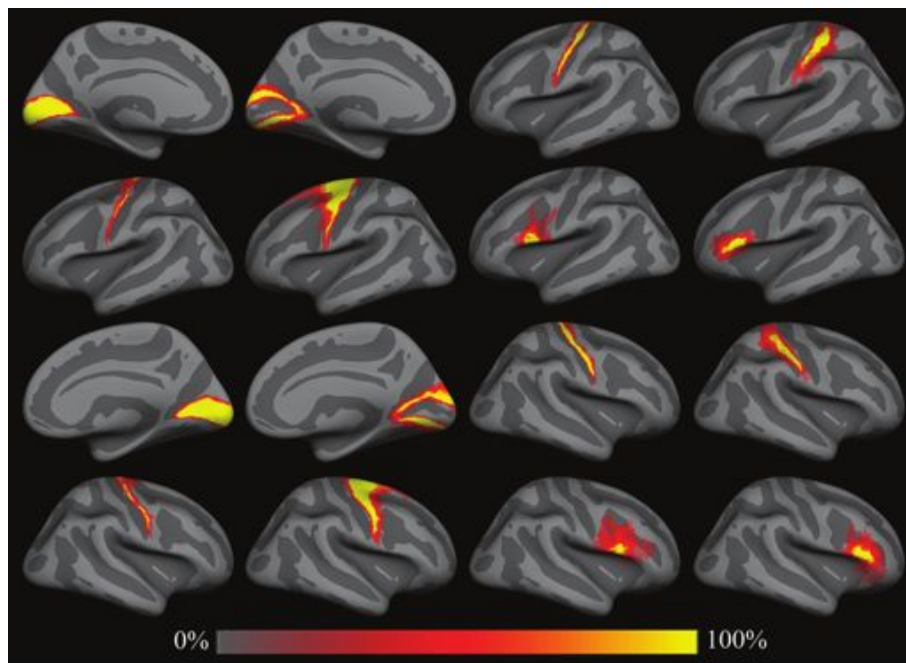


FIG. 1.18 – Cartes de probabilité surfaciques de différentes aires de Brodmann intégrées dans Freesurfer. Première ligne : aires 17, 18, 4p et 2, Deuxième ligne : aires 4a, 6, 44 et 45. Les deux dernières lignes représentent les mêmes aires que les premières mais pour l'hémisphère droit. (source : Fischl et al. (2007)).

issus de l'IRM de diffusion. Bürgel et al. (2006) fournit de cette manière des cartes probabilistes à partir de tracts identifiés sur une série de sections histologiques myélinisées pour 10 individus.

D'après l'étude précédente, les tracts se myélinisant plus tôt durant l'ontogenèse comme le tract cortico-spinal, le corps calleux, ou le tract optique ont un taux de recouvrement important à travers les sujets. A l'opposé, les tracts d'association (faisceau longitudinal supérieur, l'occipito-frontal supérieur et inférieur se myélinisent relativement tard (Yakovlev et Lecours 1967). La variabilité inter-individuelle de ces tracts est plus importante que celle des précédents. Ainsi **il semblerait que les tracts se formant plus tôt seraient les moins variables**. Bürgel donne cependant un contre-exemple avec la radiation acoustique. Etant donné le nombre de sujets (10) de l'étude, il ne faut cependant pas en tirer de

conclusions trop hâtives.

Récemment, De Schotten et al. (2011) a étudié la variabilité inter-individuelle des faisceaux de la matière blanche in-vivo à partir de tracts reconstruits à partir de données IRM de diffusion (après une modélisation des données par des tenseurs de diffusion et une tractographie déterministe, se reporter à la Partie 2.3.3 pour la définition de ces termes). Cette étude construit à partir de 40 sujets des cartes probabilistes (avec le pourcentage de recouvrement entre les sujets) des principaux grands faisceaux dans un repère de référence (MNI). La figure 1.19 montre les cartes obtenues ainsi que leur comparaison avec les cartes de l'étude histologique de Bürgel et al. (2006). Les deux études montrent une cohérence mais aussi des différences pour le tract cortico-spinal par exemple, qui est mal retrouvé par la tractographie à partir du tenseur de diffusion. Une modélisation plus fine des données de diffusion ainsi qu'une tractographie probabiliste améliorerait sa reconstruction. Rappelons que l'absence d'une vérité terrain en terme de connectivité anatomique rend difficile ce genre d'études.

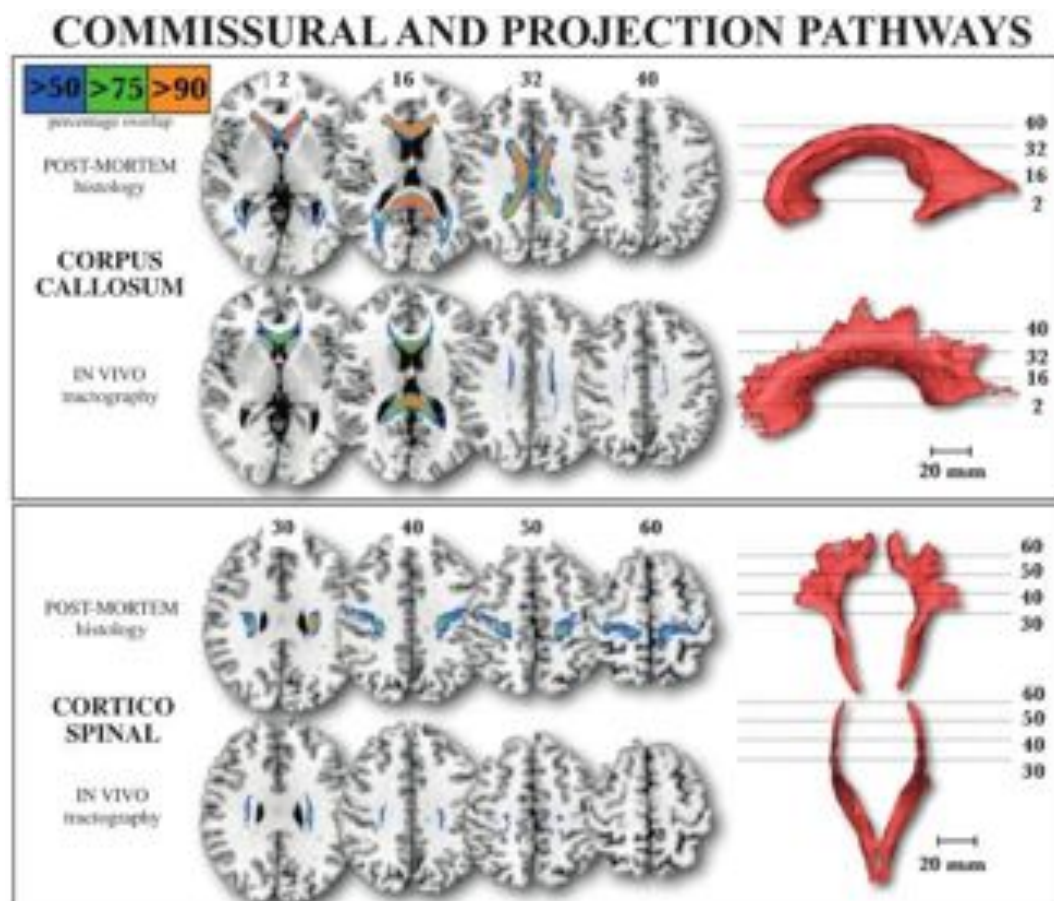
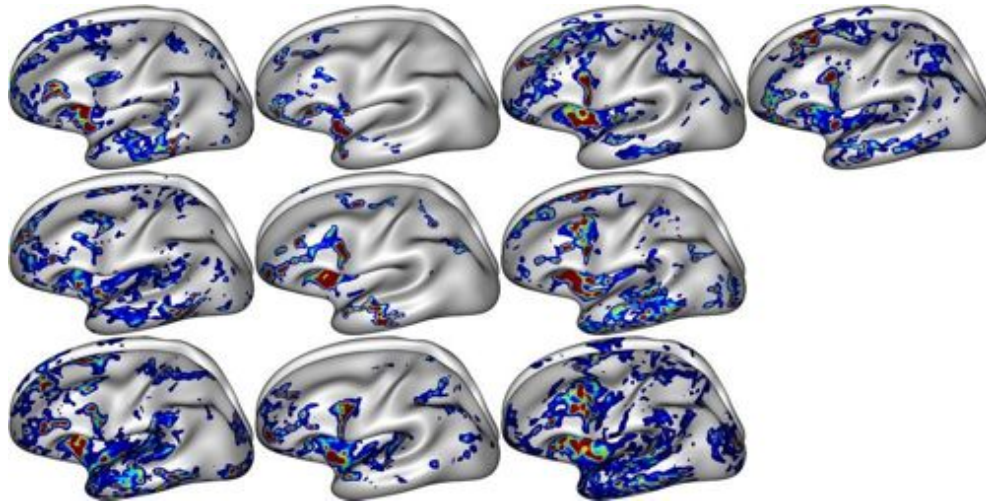
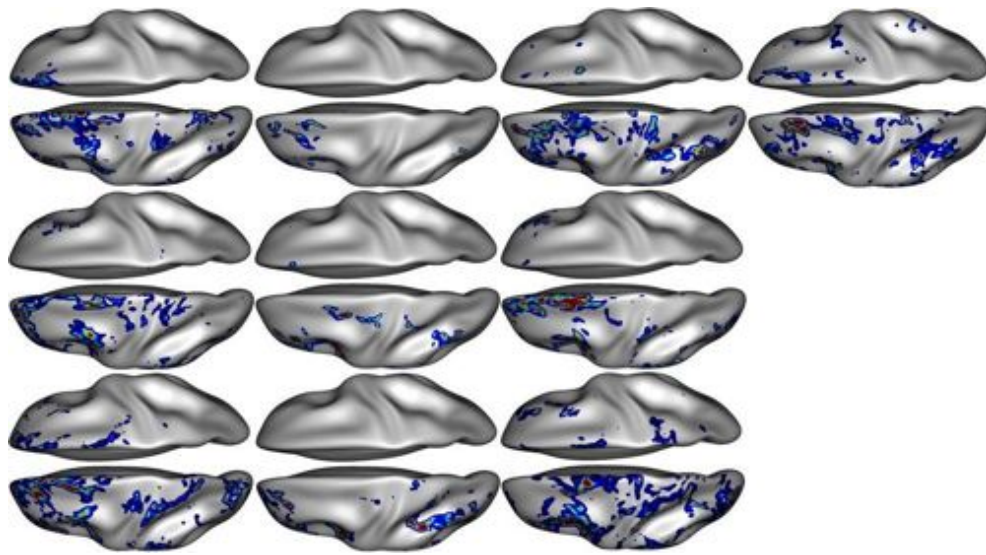


FIG. 1.19 – Comparaison des cartes de pourcentage de recouvrement issues de l'histologie (première ligne, Bürgel et al. (2006)) et issues de la tractographie (ligne du dessous) du corps calleux et du faisceau cortico-spinal. A droite, sont présentes les reconstructions 3D des cartes (pour > 50% de recouvrement entre les sujets). Image et légende tirées de De Schotten et al. (2011)

La figure 1.20 montre les cartes des connexions projetées sur la surface obtenues par tractographie sur la base de sujets étudiés pendant ma thèse pour l'aire de Broca (pour la définition de la région d'intérêt se référer à la partie 2.3.2.3).



(a) Vue de côté



(b) Vue de dessus

FIG. 1.20 – Exemple des connexions de l'aire de Broca pour dix individus (obtenues à partir d'un tracking déterministe).

Les connexions sont représentées sur un maillage moyen gonflé fourni par un traitement de Freesurfer. Ceci permet de placer les connexions de chaque sujet sur une même surface et de regarder la variabilité propre aux connexions, un recalage surfacique ayant été effectué préalablement (cf. partie 2.3.2.2 pour plus de détails). Même en s'affranchissant en partie de la variabilité de l'anatomie sulco-gyrale de cette manière, il reste une variabilité importante au niveau des connexions.

1.5 CONCLUSION

Dans ce chapitre, nous avons mentionné différentes définitions possibles du connectome cérébral humain selon l'échelle spatiale choisie ainsi que les techniques utilisées. Ce travail nous a permis de nous rendre compte de la complexité relative à la construction de ces connectomes et de la subjectivité de chaque méthode de construction.

Dans le cadre de cette thèse, **nous chercherons à identifier les noeuds d'un connectome à une échelle mésoscopique** à partir des données de connectivité issues de l'IRM de diffusion. Dans ce contexte, le type de connectome étudié est un **connectome structurel** construit à partir de données structurelles fournies **par deux modalités d'imagerie par résonance magnétique** : des images **anatomiques** ainsi que de **diffusion**.

Dans ce contexte, il est important de rappeler le **caractère approximatif des faisceaux de fibres** reconstruits à partir de l'IRM de diffusion par rapport aux fibres réelles. Ces approximations numériques seront appelées par la suite des **tracts** pour éviter toute confusion. Cette technique permet d'accéder à une approximation des faisceaux de fibres du cerveau. Ainsi, toutes les conclusions que nous pourrions tirer à partir de ces données sont à nuancer en prenant en compte cette considération.

Nous avons vu dans ce chapitre qu'il existe **dans la littérature essentiellement deux types d'approches** pour effectuer une parcellisation du cortex en fonction des données de connectivité anatomique. D'un côté une **approche aveugle** tenant compte des données de connectivité complètes vers l'ensemble du cerveau et de l'autre une **approche par régions d'intérêt a priori** utilisant par exemple une segmentation a priori du cerveau pour résumer et réduire l'information de connectivité.

Dans cette thèse **nous proposons une approche intermédiaire** entre les deux précédentes qui **identifie un ensemble de régions d'intérêt à partir des données de connectivité elles-mêmes en s'affranchissant ainsi de l'utilisation d'une segmentation a priori** et permettant une réduction de dimension des données de connectivité par rapport à l'approche aveugle (Chapitre 5). Cette approche servira ensuite pour parcelliser le cortex à l'échelle d'un sujet (Chapitre 6) ou à l'échelle d'un groupe de sujets (Chapitre 7). Nous étudierons en effet la possibilité d'effectuer cette réduction de dimension des données au niveau d'un groupe de sujets, en tenant compte de la variabilité inter-individuelle qui comme nous l'avons vu dans ce chapitre est multiple.

Avant d'aborder ces questions, nous allons présenter, dans les 3 chapitres qui suivent, les techniques d'acquisition ainsi que les outils que nous avons utilisés pour construire les connectomes structurels associés à la base de 10 sujets sains que nous étudierons pendant cette thèse.

Nous rappellerons au chapitre 2 le type de connectome que nous étudierons et les raisons qui nous ont amenés à ce choix. Nous aborderons ensuite les principes d'acquisition des images et les prétraitements effectués pour reconstruire les faisceaux de tracts et gérer la variabilité inter-individuelle des données avant d'aborder les outils que j'ai développés dans les premiers chapitres de la partie II des Contributions.

QUELS OUTILS POUR ACQUÉRIR QUEL CONNECTOME ET COMMENT ?

UN connectome structurel et surfacique, ou encore anatomique cortical

Le connectome étudié dans le cadre de cette thèse aura :

1. comme support le cortex, modélisé de manière surfacique par la **surface corticale**, extraite à partir d'images IRM anatomiques ;
2. comme mesure de connectivité : les **tracts** reconstruits à partir de l'IRM de diffusion, représentant la **structure des fibres de la substance blanche**. Cette mesure de connectivité est qualifiée d'**anatomique** (différente de la connectivité fonctionnelle obtenue par IRM fonctionnelle) ;
3. comme noeuds ou éléments de base : pas d'a priori, le but est de les identifier à partir de la **matrice de connectivité corticale** résumant les données précédentes.

Dans ce chapitre, nous justifierons dans une première partie notre choix de ce connectome structurel surfacique, motivé notamment par l'existence d'une base de données appropriée. Dans une deuxième partie, nous décrirons le principe de **l'IRM et les processus d'acquisition et de reconstruction du cortex et des tracts**. Nous décrirons ensuite comment gérer la variabilité du support de ce connectome, à travers la **mise en correspondance inter-sujets** de la surface corticale.

SOMMAIRE

3.1	INTRODUCTION	63
3.2	PARCELLISATIONS ALÉATOIRES DU CORTEX	65
3.2.1	Diagramme de Voronoï centroïdal surfacique	65
3.2.2	Applications	66
3.3	MATRICE DE CONNECTIVITÉ CORTICALE	67
3.3.1	Différentes définitions	67
3.3.2	Construction de la matrice	68
3.3.3	Connectome entre parcelles et pour un groupe de sujets .	68
3.4	VISUALISATION DE LA MATRICE DE CONNECTIVITÉ	69
3.4.1	Outils existants	69
3.4.2	Contributions personnelles	71
3.5	CONCLUSION	72

2.1 UNE ANALYSE STRUCTURELLE SURFACIQUE

2.1.1 Avantages de l'approche surfacique

Différentes raisons ont été prises en compte pour justifier le choix de l'analyse surfacique pour étudier le connectome. Voici une liste non exhaustive des principales motivations, qui reflètent les avantages de l'approche surfacique par rapport à l'approche volumique.

1. **Réduction de dimension** : la matrice de connectivité (points de la surface corticale \leftrightarrow points de la surface corticale) est réduite par rapport aux matrices de connectivité (voxels du cortex \leftrightarrow voxels de la matière blanche) étudiées dans la littérature en vue de la parcellisation du cortex ;
2. **Manque d'information issue de l'IRM de diffusion dans le cortex** : le tracking jusqu'au cortex est possible mais difficile à l'intérieur (car les modèles représentant la diffusion sont alors isotropes), l'interface cortex/matière blanche est donc plus appropriée pour étudier les extrémités de chaque tract
3. **Respect de la topologie bidimensionnelle du cortex** :
 - Prise en compte de la morphologie sulco-gyrale, importante pour éviter la confusion des fibres en U connectant deux gyri voisins
 - Métrique adaptée à la géométrie du cortex :
 La distance géodésique le long la surface corticale est plus pertinente qu'une distance euclidienne volumique, car plus fidèle à la géométrie du cortex. En effet, la distance euclidienne volumique peut sous-estimer la distance entre deux points du cortex considérés comme voisins dans le volume mais étant situés de part et d'autre d'un même sillon
 - Cohérence avec l'organisation topographique des aires sensorielles primaires : rétinotopie (vision), tonotopie (audition), somatotopie (motricité)), formant des cartes en deux dimensions
 - Cohérence avec la théorie des colonnes corticales (Mountcastle 1997), selon laquelle le cortex est organisé en colonnes (groupes de neurones) perpendiculaires à la surface corticale constituant chacune un module partageant des connexions longue distance avec les autres modules : assimiler ces colonnes aux noeuds du connectome a donc un sens selon cette théorie
4. **Existence d'outils de mise en correspondance surfacique** inter-sujets respectant la morphologie sulco-gyrale
 - Possibilité d'utiliser un **système de coordonnées surfacique** (sphérique)
 - Pas d'hypothèse d'un recalage *voxel à voxel* parfait
 - Proposition dans cette thèse d'une solution rendue possible par l'approche surfacique évitant l'hypothèse d'une correspondance parfaite *point à point* à travers les sujets
5. **Visualisation facilitée** car en deux dimensions, avec la possibilité de déplier le cortex en le gonflant artificiellement.

2.1.2 Une base de données adaptée à l'étude structurelle

En ce qui concerne la mesure de connectivité associée au cortex, le sujet de la thèse en lui-même définissait la méthode choisie. Le but de cette thèse est en effet d'étudier le connectome et d'identifier ses éléments de base (ou noeuds) à partir de la connectivité anatomique provenant de la reconstruction du réseau de fibres de la substance blanche, rendue possible in-vivo par IRM de diffusion.

Pour cela, nous disposons d'une base de donnée très appropriée car acquise dans ce but, la base NMR Poupon et al. (2006).

En effet, les **images pondérées en diffusion** de cette base ont été acquises à partir d'un imageur 1,5 T avec une **haute résolution angulaire**, dite HARDI pour *High Angular Resolution Diffusion Imaging* avec 200 directions et une *b-value* de 3000s/mm^2 . Cette résolution combinée à l'utilisation de modèles adaptés (plus fins que le tenseur de diffusion) est importante pour espérer résoudre les problèmes de croisements de fibres, très présents au voisinage du cortex. Un **ensemble de corrections** est de plus appliqué aux images. En effet, une carte de phase a également été acquise pour corriger les effets de susceptibilité magnétique. De plus, une correction de la non linéarité des gradients est effectuée à l'aide d'une représentation en harmoniques sphériques du profil du gradient. Enfin, une correction des distorsions est effectuée.

Les **images anatomiques** ont une **résolution fine** avec une épaisseur de coupe de $1,2\text{ mm}$ et une résolution dans le plan de $0,9375\text{ mm} \times 0,9375\text{ mm}$. L'alignement entre les images anatomiques et de diffusion est effectué par un recalage rigide, dont la transformation est obtenue par la maximisation de l'information mutuelle entre l'image anatomique pondérée en T1 et l'image moyenne pondérée en T2.

Cette base possède aussi des **données fonctionnelles** qui ont été à l'origine d'une parcellisation du cortex dans le cadre du travail de thèse d'Alan Tucholka (Tucholka 2010). Une comparaison avec les résultats obtenus par cette analyse et ceux obtenus par notre approche pourra être ainsi très enrichissante.

Vocabulaire

L'interface entre la matière grise et la matière blanche déterminée à partir de l'imagerie anatomique est appelée **surface corticale**. Elle sera représentée par un maillage triangulaire, i.e. par un ensemble de triangles juxtaposés dont les sommets seront appelés les **points** du maillage (Fig.2.1(a)).

Les **tracts** sont les "fibres numériques" reconstruites par tractographie. Chaque tract représente un faisceau de neurones. Par la suite, nous emploierons ce terme plutôt que celui de **fibres** pour bien différencier les fibres réelles de leur approximation numérique (cf. Fig.2.1(b) et partie 2.3.3.3 sur les algorithmes de tractographie).

La partie suivante décrit les principes physiques permettant l'acquisition de ces données ainsi que les post-traitements associés dans le contexte de cette thèse. Nous allons ainsi décrire de manière très résumée les bases de l'imagerie par résonance magnétique. Tout d'abord, nous commencerons par l'obtention des *images anatomiques* avec comme post-traitement associé l'extraction du cortex, support des noeuds du connectome. Il sera

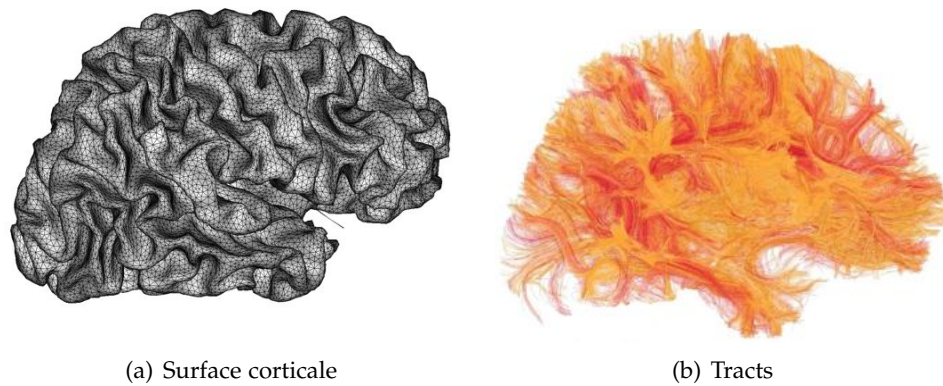


FIG. 2.1 – Les données dont la construction est expliquée dans ce chapitre : (a) Surface corticale (interface gris/blanc) d'un sujet de la base NMR, représentée par un maillage triangulaire ; (b) Exemple de tracts issus d'un tracking déterministe (cf. 2.4.1) de longueur entre 5 cm et 10 cm.

ensuite question des *images pondérées en diffusion* avec ensuite les *modèles* et les principaux algorithmes de reconstruction des tracts ou encore *tractographie*.

2.2 ACQUISITION DES IMAGES IRM

L'imagerie par résonance magnétique ou IRM est une technique non invasive permettant d'accéder à la structure, la fonction et le réseau de connectivité cérébral sous-jacent, à partir de différentes modalités : IRM anatomique, fonctionnelle, ou de diffusion. Nous aborderons ici deux modalités qui nous permettront d'extraire la surface corticale (via l'IRM anatomique) et le réseau de connectivité de la substance blanche (IRM de diffusion). Nous aurons ensuite les données pour construire un *connectome surfacique structurel* avec comme support le cortex, et comme connectivité les tracts, modélisation du réseau structurel formé par les faisceaux de neurones de la substance blanche.

2.2.1 Le phénomène de Résonance Magnétique Nucléaire

Le phénomène de *Résonance Magnétique Nucléaire* (RMN) a été découvert par Bloch et Purcell en 1946. Il concerne l'étude des modifications d'aimantation des noyaux d'une substance sous l'action conjointe de deux champs magnétiques : un *champ magnétique statique* fixe élevé \vec{B}_0 et un *champ magnétique tournant* ou *onde radio-fréquence* \vec{B}_1 . Il peut se diviser en trois étapes :

1. **Etat d'équilibre initial** : En présence d'un champ magnétique fixe \vec{B}_0 , les noyaux d'hydrogène possèdent un moment magnétique non nul, qui va s'orienter selon \vec{B}_0 (l'axe O_z), créant ainsi une aimantation macroscopique longitudinale M_z . Ils ne sont en réalité pas parfaitement alignés mais tournent individuellement avec un angle donné autour de \vec{B}_0 , avec un mouvement de précession autour de \vec{B}_0 de

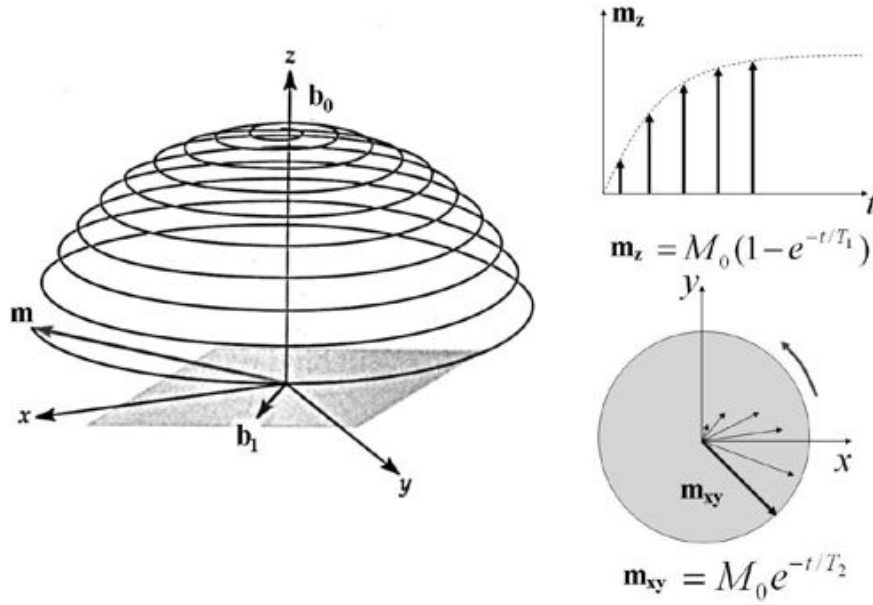


FIG. 2.2 – Principe de la relaxation des spins. Après application de l'onde excitatrice l'aimantation macroscopique \mathbf{m} suit un mouvement hélicoïdal. La repousse de l'aimantation longitudinale m_z est caractérisée par la constante de temps T_1 , tandis que l'aimantation transverse disparaît suivant la constante T_2 . (source : Huettel et al. (2004))

fréquence angulaire :

$$\omega_0 = \gamma \|\vec{B}_0\| = 2\pi\nu_0 \quad (2.1)$$

où γ est une constante appelée facteur gyro-magnétique, spécifique du noyau étudié, et ν_0 la fréquence correspondante.

Les moments magnétiques n'étant pas en phase, l'aimantation macroscopique transversale M_{xy} est nulle.

2. **Excitation** : L'application d'un champ magnétique tournant \vec{B}_1 à la fréquence ν_0 ou *fréquence de résonance* va perturber le système et faire basculer l'aimantation macroscopique dans le plan transversal avec une disparition de l'aimantation longitudinale M_z et une apparition de l'aimantation transversale M_{xy} .
3. **Retour à l'équilibre et Mesure** : Après l'arrêt de cette excitation, la dynamique de retour à l'équilibre crée une onde électro-magnétique mesurable par une antenne. La nature de cette dynamique dépend des propriétés locales des tissus (magnétiques, moléculaires...), ce qui va nous permettre de reconstruire une image contrastée.

2.2.2 Encodage spatial des images

Lors de l'application de l'onde radio-fréquence \vec{B}_1 , le système de spins entre en résonance si et seulement si la fréquence de \vec{B}_1 est égale à la fréquence de résonance du système ν_0 . En appliquant un gradient de champ $g_z = \partial B_0 / \partial z$ selon l'axe O_z , on peut modifier la fréquence de précession des spins, qui dépendra maintenant de la position suivant z . Ainsi, lors de

Tissus	T_1 (ms)	T_2 (ms)
matière grise	900	100
matière blanche	600	80
Liquide céphalo-rachidien	4 000	2 000

FIG. 2.3 – Temps de relaxation des principaux tissus cérébraux à 1,5 Tesla. (source : Huettel et al. (2004))

l'excitation du système par une onde radio-fréquence excitatrice \vec{B}_1 centrée autour d'une fréquence ν (avec une largeur $\Delta\nu$), celle-ci n'interagira qu'avec les spins qui tournent à cette fréquence, correspondant ainsi à une coupe de largeur $\Delta\nu$ suivant l'axe 0_z . On parle alors d'excitation *sélective*. L'encodage dans le plan transversal s'effectue d'une manière analogue avec l'application d'un gradient de codage de phase et d'un gradient de fréquence. L'image est ensuite reconstruite en effectuant une transformée de Fourier inverse (cf. Kastler et al. (2006) pour plus de détails).

2.2.3 Imagerie anatomique : Contrastes en T_1 et T_2

Lors du phénomène de résonance magnétique nucléaire, le retour à l'équilibre se traduit par deux types de *relaxations* faisant intervenir des mécanismes différents (cf. Fig. 2.2) :

- la *repousse de l'aimantation longitudinale* M_z , reposant sur les échanges énergétiques des spins avec le milieu environnant (interaction spin-milieu), avec une constante de temps T_1 ,
- la *disparition progressive de l'aimantation transversale* M_{xy} , résultant d'un déphasage des spins dans leur mouvement de précession (interaction spin-spin), avec une constante de temps T_2 .

Ces deux constantes de temps dépendent des propriétés des tissus cérébraux. Cette hétérogénéité va permettre de distinguer les différents tissus et de reconstruire une image contrastée (cf. Fig 2.3). De manière très simplifiée (cf. Kastler et al. (2006) pour plus de détails), une **séquence d'acquisition** est constituée d'une suite d'opérations (excitation par une onde radio-fréquence, application de gradients) qui vont permettre de pondérer l'image obtenue par rapport à l'une ou l'autre de ces constantes de temps afin d'obtenir les contrastes associés.

2.2.4 IRM de diffusion : pour accéder aux tracts

L'imagerie par IRM de diffusion est **sensible au mouvement des molécules d'eau** dans le cerveau (Le Bihan et al. 1986). C'est à ce jour la seule technique **non invasive** qui permet de sonder la **structure de la matière blanche**.

2.2.4.1 Phénomène de diffusion

A l'échelle microscopique les molécules ont un mouvement aléatoire ou **mouvement brownien** dû à l'agitation thermique. Ce phénomène de diffusion peut être modélisé par la relation d'Einstein (1905) selon laquelle la variance $\langle R^2 \rangle$ du déplacement des molécules en trois dimensions dans

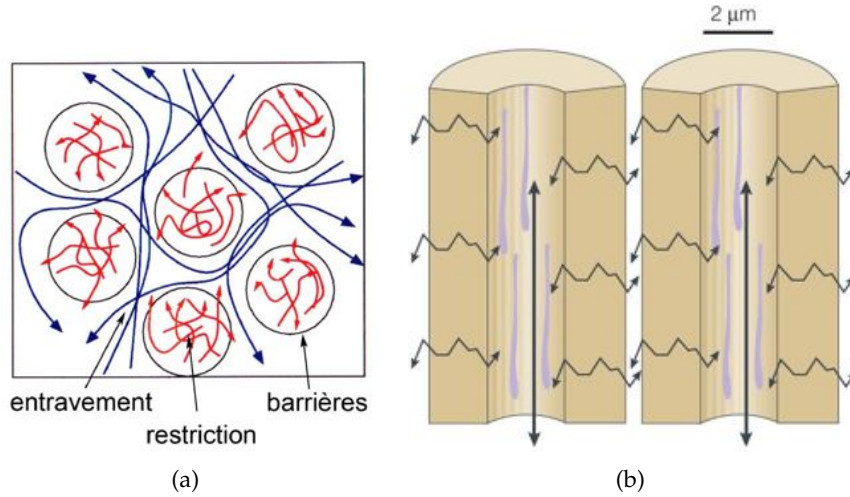


FIG. 2.4 – Diffusion anisotrope dans la matière blanche. 2.4(b) Dans la direction des fibres (D_{\parallel}) la diffusion est plus importante que dans la direction perpendiculaire (D_{\perp}) où la diffusion est entravée. $D_{\parallel} = 1,2 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$ et $D_{\perp} = 0,4 \cdot 10^{-3} \text{ mm}^2\text{s}^{-1}$ (source : Le Bihan (2003)); 2.4(a) La diffusion peut être considérée comme restreinte dans l'espace intra-cellulaire, entravée dans l'espace extra-cellulaire et modifiée par l'échange au travers de barrières semi-perméables entre les deux espaces (source : Poupon (1999)).

un milieu isotrope est une fonction linéaire du temps :

$$D = \frac{1}{6\tau} \langle R^T R \rangle \text{ avec : } \begin{cases} D : \text{coefficient de diffusion} \\ R : \text{vecteur de déplacement} \\ \tau : \text{temps de diffusion.} \end{cases} \quad (2.2)$$

Lorsque le milieu est anisotrope (comme c'est le cas des tissus cérébraux), le coefficient de diffusion scalaire D de la relation précédente est remplacé par un **tenseur de diffusion** :

$$\mathbf{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix}$$

Le phénomène de diffusion peut être considéré comme **restreint dans l'espace intra-cellulaire**, et **entravé dans l'espace extra-cellulaire** (cf. Fig. 2.4(a)). Dans la matière blanche, la diffusion est plus importante dans la direction des faisceaux de fibres que perpendiculairement : le mouvement des molécules d'eau est principalement orienté le long des faisceaux de fibres (cf. Fig. 2.4(b)).

2.2.4.2 Imagerie et Séquences d'acquisition

Stejskal et Tanner (1965) proposent une séquence d'imagerie pondérée en diffusion appelée Pulse Gradient Spin Echo (PGSE). L'idée est d'effectuer un **marquage en phase des spins selon leur position** dans la coupe acquise.

Ceci est effectué par l'application après l'onde excitatrice radio-fréquence (RF) à 90° d'un gradient de champ magnétique dit de diffusion \mathbf{G}_D (cf. Fig. 2.5). Ce gradient va introduire un déphasage des spins lors de

leur retour à l'équilibre. L'application successive d'une onde RF à 180° qui inverse les déphasages, puis d'un gradient de diffusion identique au précédent va effacer le marquage en phase des spins introduit par le premier gradient : il y a rephasage des spins.

Les **spins diffusants** qui se sont déplacés durant cette séquence n'auront pas subi des gradients identiques et ne vont pas se rephaser exactement. Il y a alors une perte de signal qui peut s'écrire sous la forme :

$$S = S_0 e^{-T_E/T_2} e^{-bD(\mathbf{o})} \text{ avec : } \begin{cases} \mathbf{o} = (G_D^x, G_D^y, G_D^z)^T / \|\mathbf{G}_D\| \\ b = \gamma^2 G_D^2 \delta^2 (\Delta - \delta/3), \text{ le degré de} \\ \text{pondération en diffusion.} \\ D \text{ le coefficient de diffusion apparent} \\ \text{(mesuré)} \end{cases} \quad (2.3)$$

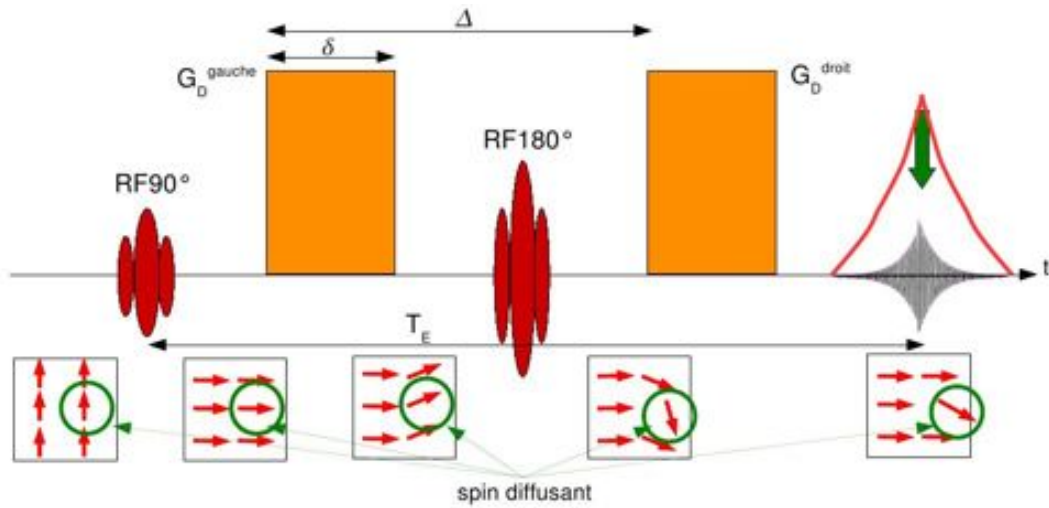


FIG. 2.5 – Séquence Pulse Gradient Spin Echo (PGSE) de Stejskal et Tanner (1965). Les spins sont marqués en phase par un premier gradient de diffusion \mathbf{G}_D . Après inversion de leur phase par impulsion radio-fréquence à 180° , et application d'un second gradient identique au premier \mathbf{G}_D , il y a rephasage des spins statiques. Les spins diffusants ne sont pas rephasés et entraînent donc une perte de signal ou hyposignal.

La diffusion perpendiculairement au gradient \mathbf{G}_D n'empêche pas les spins de se rephaser. Ainsi lors de l'application d'un double gradient de diffusion \mathbf{G}_D , le signal est sensible à la diffusion dans la direction ou **orientation** \mathbf{o} du gradient.

Le **degré de pondération en diffusion** b appelé également **facteur b** (Le Bihan et al. 1986, Basser et al. 1994) règle comme son nom l'indique la pondération en diffusion appliquée et dépend des paramètres d'acquisition, tandis que D dépend du milieu de diffusion. Ainsi :

- si $b \ll D$: pour toutes les directions d'application du gradient, l'atténuation du signal est faible, le profil de diffusion obtenu est alors sphérique (cas où le milieu est isotrope) ;
- si $b \gg D$: il y a une bonne différenciation entre les orientations présentant une diffusion élevée et celles de basse diffusivité, le profil de diffusion obtenu est plus précis mais le rapport signal sur bruit diminue.

2.3 PRÉTRAITEMENTS DES IMAGES IRM

Dans cette partie, nous commencerons par présenter les différents logiciels dont nous nous sommes servis pendant cette thèse. Nous décrirons ensuite plus en détails les prétraitements des données anatomiques puis ceux des données de diffusion en privilégiant les méthodes que nous avons utilisées.

2.3.1 Les logiciels utilisés

Brainvisa - Traitements anatomiques et visualisation

Brainvisa¹ est un logiciel dédié au traitement d'images cérébrales développé principalement au Laboratoire de Neuro-imagerie par Ordinateur de Neurospin (CEA). L'une de ses spécialisations principales est une chaîne de traitements des images anatomiques allant de la segmentation du cortex jusqu'à l'identification automatique des sillons corticaux (Rivière et al. 2002), en passant par une reconstruction du maillage de la surface corticale (Mangin et al. 1995). Ces traitements seront décrits dans la partie 2.3.2.1 suivante.

Anatomist Anatomist est un logiciel de visualisation d'images 3D et de maillages (et tracts) développé par Denis Rivière au sein de Neurospin. Il est incorporé dans Brainvisa mais peut fonctionner indépendamment et est scriptable en python. La plupart des images utiliseront ce logiciel. Nous proposerons de plus un nouveau type de visualisation du connectome grâce à cet outil, au chapitre 3.

Ptk - Données de diffusion

Ptk est un outil spécialisé dans les images de diffusion initié et développé principalement par Cyril et Fabrice Poupon au sein de Neurospin (CEA). Il sera bientôt incorporé dans Brainvisa pour former la toolbox Connectomist dédiée aux images de diffusion. Nous présenterons dans la partie 2.3.3 certains traitements dont la modélisation locale des données de diffusion (Descoteaux et al. 2007) ainsi que des algorithmes de tractographie (Perrin et al. 2008).

Freesurfer - Traitements anatomiques inter-sujets

Freesurfer² est un logiciel de traitement d'images cérébrales spécialisé dans l'analyse de la surface corticale et développé au Martinos Center for Biomedical Imaging. Il fournit notamment un ensemble d'outils permettant de reconstruire les maillages de la surface corticale (Dale et al. 1999) et de la parcelliser en gyri (Desikan et al. 2006). Une analyse surfacique intra et inter-sujets permet de plus d'associer au cortex un système de coordonnées sphériques qui sera utile pour les analyses de groupe (Fischl et al. (1999b;a)). Nous détaillerons ces étapes dans la partie 2.3.2.2.

¹<http://brainvisa.info>

²<http://surfer.nmr.mgh.harvard.edu>

2.3.2 Prétraitements anatomiques

Le traitement anatomique de la base de données utilisée pendant cette thèse résulte de deux logiciels essentiellement : Brainvisa ³, pour l'extraction de la surface corticale (intra-sujet) et Freesurfer ⁴fournissant un système d'extraction et de mise en correspondance inter-sujets des surfaces corticales. Ces traitements anatomiques ont été effectués par Alan Tucholka au début de sa thèse. Nous n'allons pas décrire ici en détail ces traitements puisqu'il ne font pas partie intégrante de mon travail. Une description est cependant nécessaire pour comprendre les données que j'ai eues à ma disposition. Le lecteur intéressé pourra se référer à la thèse d'Alan Tucholka Tucholka (2010). Ces données anatomiques seront ensuite combinées aux données de diffusion grâce à des outils que j'ai développés pendant ma thèse (cf. partie 3).

A titre d'exemple, nous décrirons brièvement les prétraitements de Brainvisa aboutissant à l'extraction de la surface corticale. Nous présenterons ensuite la méthode de mise en correspondance surfacique inter-sujets de Freesurfer. Dans un dernier point nous parlerons des régions d'intérêt qui seront ensuite utilisées (comme régions à parcelliser). Nous décrirons ainsi comment j'ai défini une région englobant l'aire motrice supplémentaire (notée SMA pour Supplementary Motor Area) et l'aire supplémentaire pré-motrice (ou pré-SMA).

2.3.2.1 Extraction de la surface corticale : exemple de Brainvisa

Les différentes étapes de la chaîne de traitements utilisées dans le cadre de cette thèse sont les suivantes (cf. Fig. 2.6) :

1. Placement dans le référentiel de Talairach par l'intermédiaire du positionnement manuel des commissures antérieure et postérieure (CA/CP) ;
2. **Correction du Biais** : Les images IRM présentent un biais spatial en intensité résultant de l'inhomogénéité du champ magnétique : le même tissu peut avoir une intensité différente suivant sa position dans l'image (cf. Fig. 2.6(a)). La correction à effectuer est estimée pour chaque sujet en minimisant une mesure de dispersion des intensités (l'entropie de leur distribution, Mangin (2000)) ;
3. **Analyse d'histogramme** des intensités : dans le but de détecter les modes de l'histogramme correspondant à la matière grise et à la matière blanche ;
4. **Segmentation du cerveau** à partir de l'étape précédente après binarisation de l'image en gardant la plage des niveaux de gris des tissus cérébraux et en effectuant différentes opérations morphologiques.
5. **Séparation des hémisphères** : Ce traitement sépare la matière blanche en cinq parties (cf. Fig. 2.6(b)) : les deux hémisphères cérébraux, le tronc cérébral, et les hémisphères du cervelet, à partir d'une série de traitements morphologiques et de la position de CA/CP et du plan interhémisphérique ;

³<http://brainvisa.info>

⁴<http://surfer.nmr.mgh.harvard.edu/>

6. **Segmentation du cortex et construction des maillages du cortex :**
 Cette procédure construit notamment une interface gris/blanc à topologie sphérique qui va être très utile et permettre notamment : le lissage géodésique sur la surface corticale (de la matrice de connectivité anatomique) ainsi que le gonflement de cette surface (très utile pour la visualisation). Une suite de transformations homotopiques d'un maillage de la boîte englobante du cerveau aboutit ainsi à un objet dont la frontière externe est l'interface Gris/LCR et la frontière interne l'interface Gris/blanc (cf. Fig. 2.6(d)).

2.3.2.2 Modélisation de la variabilité interindividuelle corticale de Freesurfer

Le logiciel FreeSurfer, développé au Martinos Center for Biomedical Imaging (Institut Harvard/MIT) propose un moyen pour mettre en correspondance les cortex d'un groupe de sujets au niveau surfacique.

Après l'extraction des surfaces corticales pour chaque sujet (Dale et al. 1999), cette approche se déroule comme suit :

1. Etape intra-sujet : Déformation de la surface corticale de chaque sujet vers une sphère (Fischl et al. 1999b) (cf. Fig. 2.7(a)) Le calcul de la transformation est effectué :
 - (a) En minimisant les distorsions de distance,
 - (b) A partir d'informations géométriques du cortex (courbure par exemple);
2. Inter-sujet : Déformation de la sphère de chaque sujet vers un modèle (Fischl et al. 1999a) (cf. Fig. 2.7(b)).

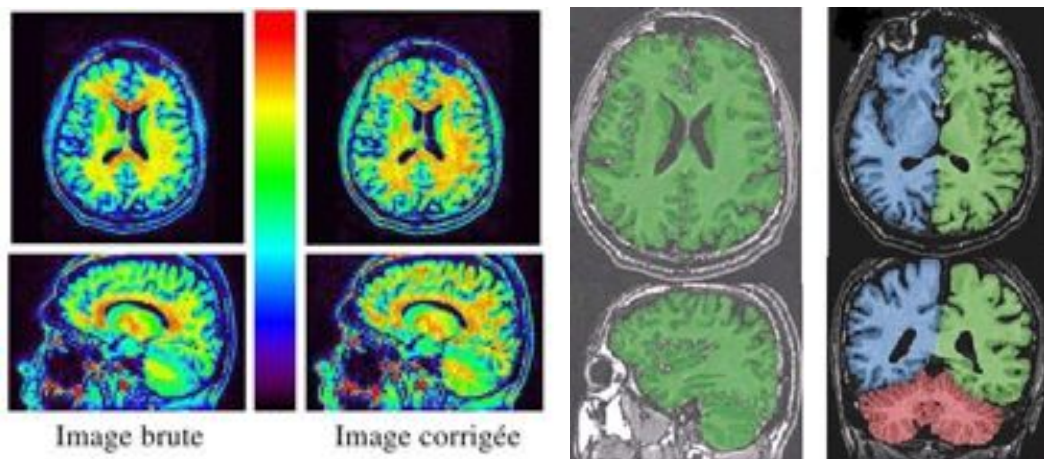
Alan Tucholka, dans son travail de thèse (Tucholka 2010), a ensuite complété cette approche par une autre étape de rééchantillonnage des points du cortex de chaque sujet, de manière à avoir le même nombre de points ainsi qu'une correspondance point à point entre les sujets.

Voici quelques propriétés à retenir de cette analyse inter-sujets :

- elle est **surfacique**,
- elle fournit des **surfaces corticales avec une correspondance point à point à travers les sujets**,
- le recalage sur le modèle s'effectue suivant la courbure du maillage donc indirectement sur les **sillons et les gyri du cortex** : ils sont en **correspondance** entre les individus
- il ne faut pas oublier que la correspondance point à point est **imparfaite**.

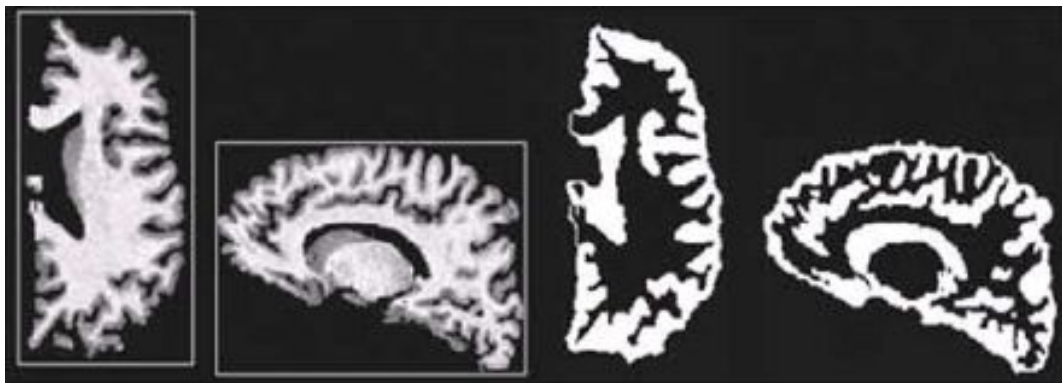
2.3.2.3 Définition de régions corticales d'intérêt

Segmentation en gyri Freesurfer fournit une segmentation en 66 régions (Desikan et al. 2006). A partir de celle-ci, nous allons pouvoir définir des régions d'intérêt : les régions définies par Freesurfer elles-mêmes ou un regroupement de ces régions. Ces régions sont définies pour l'ensemble des sujets. Cela va grandement simplifier la construction de régions d'intérêt.

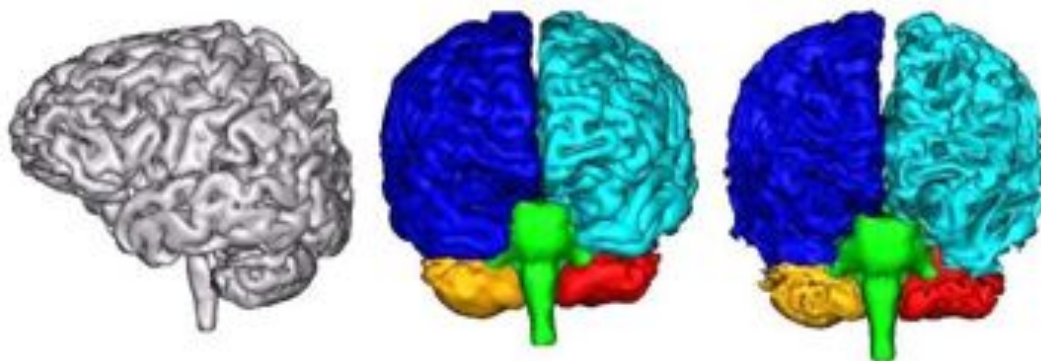


(a) Correction du biais

(b) Segmentation du cerveau : enveloppe externe, hémisphères (vert et bleu), cervelet (rouge).

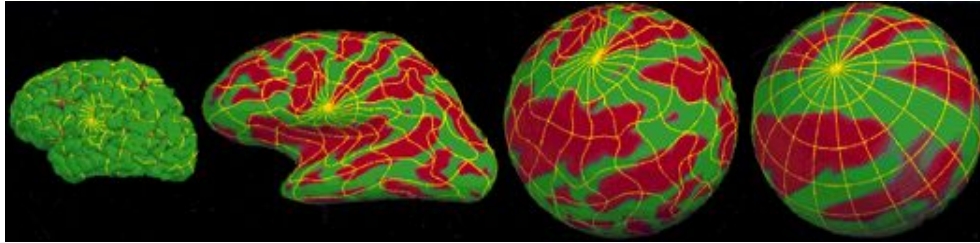


(c) Segmentation du complexe matière grise/liquide céphalo-rachidien. A gauche, images initiales, à droite images binarisées.

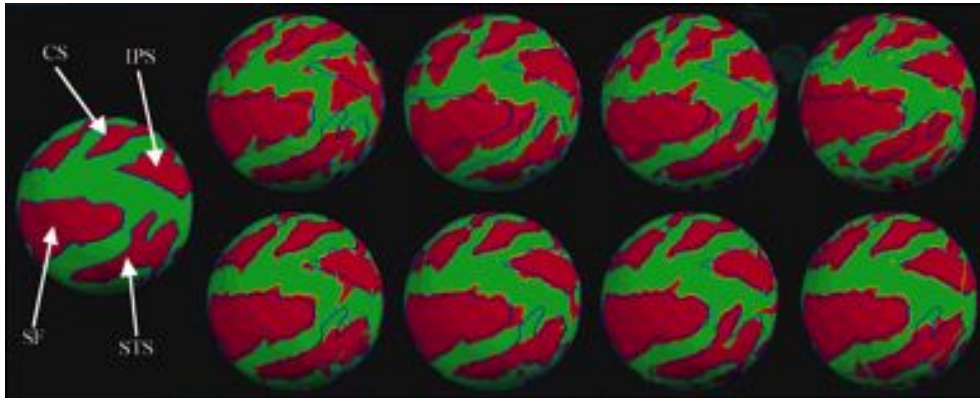


(d) Les différents maillages obtenus. A gauche et au centre : maillage représentant l'interface gris/LCR, à droite : maillages de l'interface gris/blanc.

FIG. 2.6 – Segmentation du cerveau - Les différentes étapes effectuées par le logiciel Brainvisa.



(a) Niveau individuel : Les différentes représentations de la surface corticale d'un individu, et la moyenne sur 40 sujets à l'extrême droite.



(b) Recalage sur un modèle : Quatre hémisphères individuels avant (haut) et après recalage (bas) avec la moyenne sur 40 sujets du pattern sulco-gyral. CS = sillons central, IPS = sillon intrapariétal, SF = fissure sylvienne, STS = sillons temporal supérieur.

FIG. 2.7 – Modélisation inter-sujets surfacique de Freesurfer. 2.7(a) Analyse individuelle, 2.7(b) Analyse inter-sujets. Les figures sont extraites de l'article Fischl et al. (1999a)

Région englobant le SMA et le pré-SMA A partir d'un sujet, une région englobant l'aire motrice supplémentaire (SMA) et pré-SMA est définie (d'une manière analogue à Johansen-Berg et al. (2004), Anwander et al. (2007) et Lehericy et al. (2004)). Nous appellerons dans la suite cette région $R_{SMA \cup \text{pré-SMA}}$. La délimitation antérieure est une ligne virtuelle verticale (dans l'espace de Talairach) passant par le genou du corps calleux (cf. Fig. 2.8(a)). La délimitation postérieure est le gyrus précentral, le gyrus caudal moyen (partie ventrale) et le gyrus para-central (partie médiale) tels que définis par la parcellisation de Freesurfer décrite précédemment (cf. Fig. 2.8(b) et 2.8(c)).

La région d'intérêt définie pour ce sujet est ensuite reportée sur les autres sujets par l'intermédiaire de la correspondance point à point des maillages fournie par freeSurfer (cf. partie précédente). Ceci est fait de manière à avoir pour chaque sujet une région d'intérêt avec exactement le même nombre de points, avec de plus une correspondance entre les points des différents sujets au sein de la région d'intérêt. Par la suite, le moyennage des données de connectivité point par point sera alors immédiat.

Broca La région de Broca est définie en regroupant les régions **Pars Triangularis** et **Pars Opercularis** de la parcellisation de Freesurfer.

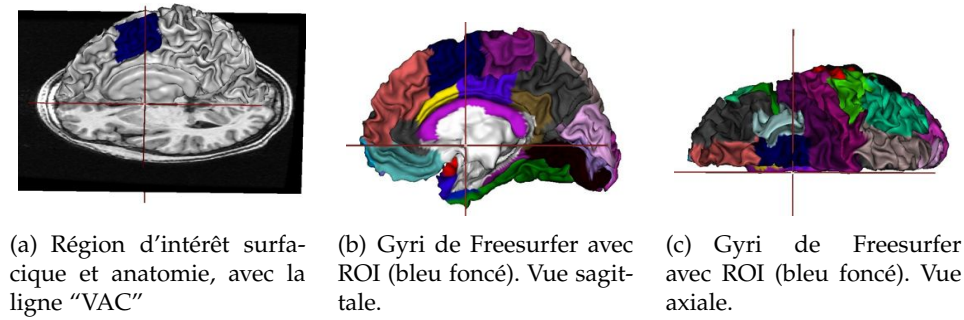


FIG. 2.8 – Région d'intérêt englobant le SMA et le pré-SMA. Sur chacune des figures, la ligne verticale correspond à la ligne verticale passant par la commissure antérieure dans le repère de Talairach ("Vertical line from the Anterior Commissure" appelée **VAC line**) qui est connue pour séparer le SMA et le pré-SMA.

2.3.3 Prétraitements des images de diffusion

Les données de diffusion de la base de données de cette thèse ont été traitées à l'aide de **Ptk** dont les outils principaux seront bientôt diffusés et incorporés dans une nouvelle toolbox de Brainvisa nommée **Connectomist**. Les deux types de traitements que j'ai utilisés sont le traitement de la modélisation locale des données de diffusion ainsi que des algorithmes de tractographie qui reconstruisent les faisceaux de fibres à partir des données locales. Dans cette partie nous présenterons les différents types de modélisation locale dont celle que nous avons choisie et pourquoi. Nous présenterons ensuite les algorithmes de tractographie déterministe et probabiliste que nous avons utilisés.

2.3.3.1 Données locales de diffusion

A partir des images IRM pondérées en diffusion, différentes techniques ont été proposées pour estimer l'orientation des fibres. Ces techniques proposent **différentes façons de modéliser différents types d'information**. La figure 2.9 présente ainsi les différents types d'information locale auxquels il est possible d'avoir accès.

Le propagateur de diffusion Callaghan et al. (1990) a défini ainsi le *propagateur de diffusion* $P(\mathbf{r}, \mathbf{r}_0, \tau)$ comme la probabilité qu'une particule soit en \mathbf{r} au temps t alors qu'elle était en \mathbf{r}_0 au temps $t - \tau$. Si l'on arrive à avoir une estimation de ce propagateur, on aura indirectement une information sur l'orientation des faisceaux de fibres. En IRM l'échelle d'observation est le voxel de l'ordre du mm^3 . Le propagateur de diffusion que l'on cherche à estimer est donc un **propagateur de diffusion moyen** que nous appellerons $P(\mathbf{r}, \tau)$ avec :

$$P(\mathbf{r}, \tau) = \int P(\mathbf{r}_1, \mathbf{r}_0, \tau) p(\mathbf{r}_0) d\mathbf{r}_0$$

où $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_0$ est le déplacement relatif de la particule et $p(\mathbf{r}_0)$ la densité initiale de particules.

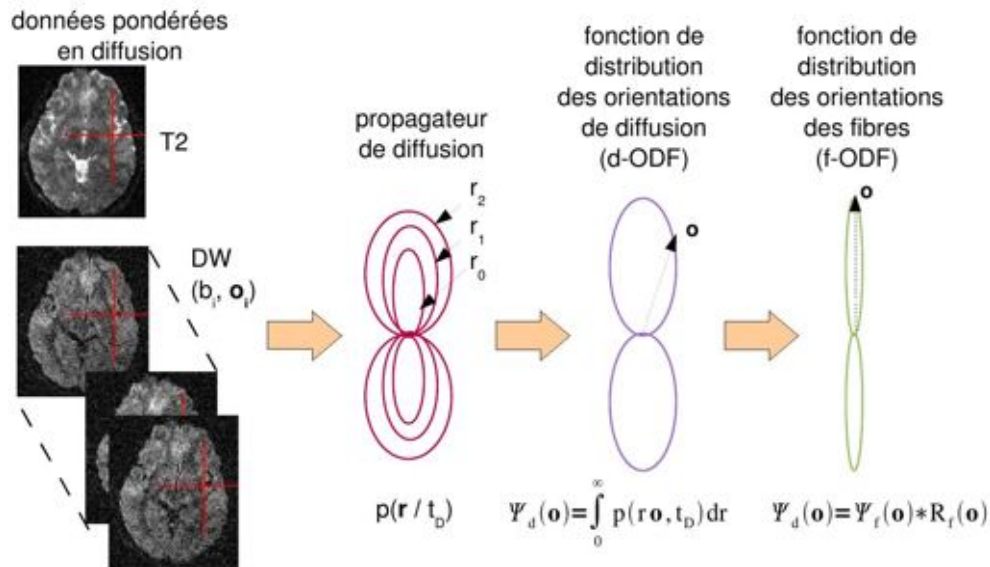


FIG. 2.9 – Données de diffusion. Les différents types d'information locale à modéliser. (source : Cours de Cyril Poupon 2010)

La fonction de distribution des orientations de diffusion ou d-ODF pour "Orientation Distribution Function" est définie comme la projection radiale du propagateur de diffusion, selon l'équation :

$$\phi(u) = \int_0^\infty P(\rho u) d\rho. \quad (2.4)$$

La fonction de distribution des orientations de fibres ou f-ODF représente la distribution des fibres comme son nom l'indique et est reliée à la d-ODF par une convolution de celle-ci avec la réponse moyenne d'un faisceau de fibres (cf. relation de la figure 2.9).

Le modèle le plus répandu est celui du **tenseur de diffusion (DTI)** utilisé couramment en clinique. Il permet de retrouver l'orientation principale des fibres de la substance blanche lorsque celles-ci sont parallèles mais ne modélise pas les croisements (accès à la d-ODF modélisée par une gaussienne).

L'**imagerie de diffusion à haute résolution angulaire (HARDI)** pour High Angular Resolution Diffusion Imaging) n'est pas applicable en clinique car elle nécessite plus de temps d'acquisition mais permet en revanche, combinée à une modélisation plus fine des informations locales de diffusion, de modéliser les croisements de fibres. Certains modèles permettent de plus de remonter jusqu'à la fonction de distribution des orientations de fibres (f-ODF). Nous allons dans la suite présenter brièvement le principe de ces différentes techniques.

2.3.3.2 Modèles locaux et autres techniques

Tenseur de diffusion Ce modèle fait l'hypothèse d'une diffusion libre et suppose qu'il n'y a pas de volume partiel, c'est-à-dire qu'il n'y a qu'une seule orientation de fibres par voxel. Il revient à une modélisation gaussienne de la fonction de distribution des orientations de diffusion d-ODF.

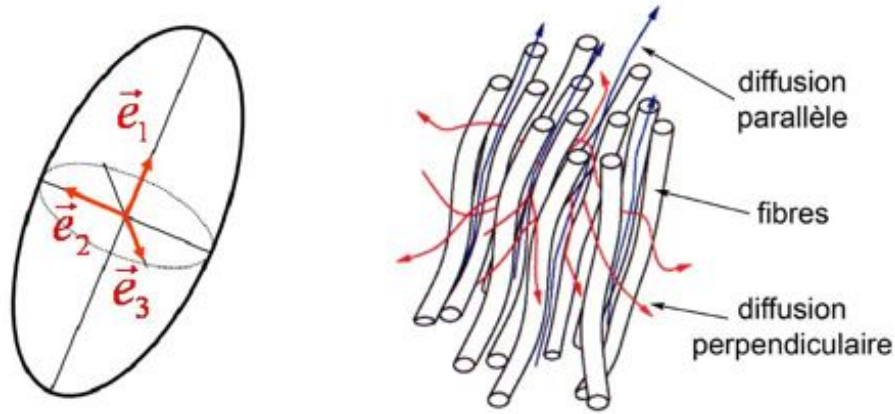


FIG. 2.10 – Modèle du tenseur de diffusion dans un milieu anisotrope. A gauche, l'axe principal de l'ellipsoïde formée par les vecteurs propres du tenseur de diffusion indique la direction privilégiée des fibres. A droite, rappel : la diffusion est plus importante le long des faisceaux de fibres. (source : Poupon (1999))

En effet, dans un milieu où la diffusion est libre et où il n'y a pas de volume partiel dans le voxel, le propagateur de diffusion moyen est gaussien et peut s'écrire :

$$P(\mathbf{r}, \tau) = \frac{1}{(4\pi\tau)^3 |\mathbf{D}|} e^{-\frac{1}{4\tau} \mathbf{r}^T \mathbf{D}^{-1} \mathbf{r}} \quad (2.5)$$

De plus, rappelons qu'en milieu anisotrope le signal mesuré en IRM de diffusion s'écrit :

$$S(\mathbf{q}, \tau) = S_0 e^{-T_E/T_2} e^{-\tau \mathbf{q}^T \mathbf{D} \mathbf{q}}.$$

Pour estimer le tenseur de diffusion \mathbf{D} il suffit alors de mesurer le signal dans six directions non parallèles.

Le tenseur est une matrice symétrique positive diagonalisable en base orthonormée. Sa diagonalisation donne 3 valeurs propres $\lambda_1, \lambda_2, \lambda_3$ et vecteurs propres $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ dont le premier vecteur propre représente la direction supposée des fibres (cf. Fig. 2.10)

Le modèle du tenseur de diffusion est avantageux car il est robuste (peu de paramètres) et utilisable en clinique. Cependant, il fait l'hypothèse d'une diffusion libre et de l'existence d'une seule direction de fibres possible par voxel. Cette dernière hypothèse rend impossible la modélisation des croisements de fibres. En effet, dans ce cas l'ellipsoïde du tenseur sera une sphère comme dans le cas d'un milieu isotrope.

D'autres façons de retrouver les directions principales des fibres sont apparues pour essayer de résoudre ce problème. Ces méthodes se divisent en deux types, soit des extensions du modèle précédent, soit des méthodes n'utilisant aucun modèle.

Extension du DTI Une première solution consiste à étendre le modèle du tenseur de diffusion en un modèle dit **multi-tenseurs** où l'on fait l'hypothèse que dans chaque voxel la diffusion est modélisable par un mélange de lois gaussiennes (Alexander et al. (2002), Tuch (2002)).

Les techniques d'imagerie de diffusion sans modèle Une deuxième option consiste à n'utiliser aucun modèle pour retrouver le profil de diffusion. C'est le cas de la **DSI** (pour Diffusion Spectrum Imaging) (Frank (2001), Tuch (2002)) ou du **Q-ball numérique** (Tuch et al. (2003), Tuch (2004)). Ces techniques reposent sur une relation existant entre le signal mesuré et le propagateur de diffusion.

En effet, en se plaçant dans l'espace \mathbf{q} , i.e. en définissant le vecteur d'onde \mathbf{q} par :

$$\mathbf{q} = \gamma\tau\mathbf{G}_D, \text{ avec } \tau = \frac{\delta}{2\pi}, \quad (2.6)$$

nous avons en milieu isotrope :

$$S(\mathbf{q}) = S_0 e^{-T_E/T_2} e^{-|\mathbf{q}|^2 \tau D}, \text{ avec } b = |\mathbf{q}|^2 \tau,$$

et en milieu anisotrope (Basser et al. 1994) :

$$S(\mathbf{q}) = S_0 e^{-T_E/T_2} e^{-\tau \mathbf{q}^T \mathbf{D} \mathbf{q}}.$$

Callaghan (1993) a alors montré que le propagateur de diffusion et le signal mesuré dans l'espace \mathbf{q} sont reliés par une transformée de Fourier :

$$\frac{S(\mathbf{q})}{S_0} = \int_{\mathbb{R}^3} p(r, \tau) e^{-i2\pi \mathbf{q}^T \mathbf{r}} d\mathbf{r} = F[p(\mathbf{r})] \quad (2.7)$$

Cette relation forme la base de l'imagerie de l'espace \mathbf{q} . De cette relation, il est possible d'obtenir le propagateur moyen par simple transformée de Fourier inverse avec :

$$p(\mathbf{r}) = F^{-1} \left[\frac{S(\mathbf{q})}{S_0} \right]. \quad (2.8)$$

Rappelons que dans le cas du modèle gaussien, cette fonction est gaussienne avec comme matrice de variance-covariance le tenseur de diffusion. La fonction de distribution des orientations de diffusion ou ODF ("Orientation Distribution Function") s'obtient ensuite par projection radiale à partir de l'équation 2.4.

DSI Dans le cas de la DSI, le calcul de la d-ODF est effectué à partir de l'équation d'inversion 2.8 et d'un échantillonnage cartésien du signal, ainsi que de la formule de projection radiale 2.4.

Q-ball numérique Dans le cas du Q-ball, le signal est échantillonné sur une sphère, et l'ODF est obtenue par une transformée de Funk, qui s'avère être une bonne approximation de la d-ODF (démonstration de Tuch (2004)).

Le Q-ball analytique Cette approche repose sur une décomposition en harmoniques sphériques HS du signal Descoteaux et al. (2007). Le théorème de Funk-Hecke permet alors de démontrer qu'il existe un lien entre la décomposition en harmoniques sphériques du signal et celle de l'ODF de diffusion. Ceci permet ainsi de reconstituer l'ODF de diffusion à partir de la décomposition HS du signal.

Cette modélisation permet d'accéder à l'ODF de fibre par déconvolution de l'ODF de diffusion avec la réponse moyenne d'un faisceau de fibres. Une compression importante de l'information est effectuée et la méthode est donc beaucoup plus rapide que celle du Q-ball numérique. L'ODF de fibre obtenue est de plus dotée d'une meilleure résolution angulaire que l'ODF de diffusion. Pour ces raisons, nous avons choisi cette méthode pour modéliser l'information locale de diffusion.

Après cette étape, un algorithme de tractographie va permettre de reconstruire une approximation de la trajectoire de chaque faisceau de fibres appelée **tract** à partir des f-ODF ou d-ODF modélisées. Nous décrivons les principales méthodes utilisées dans la suite.

2.3.3.3 Tractographie

Les algorithmes reconstruisant les faisceaux de fibres en trouvant une approximation de leur trajectoire sont appelés des algorithmes de **tractographie**. Les faisceaux reconstruits sont appelés des **tracts** et sont des approximations numériques des fibres réelles. Il existe différents types de tractographies traités de différentes façons (nous ne ferons pas une bibliographie exhaustive ici mais nous donnerons juste des exemples) :

- les tractographies déterministes par ligne de courant, initiées par Mori et al. (1999),
- les tractographies géodésiques (Lenglet (2006), Jbabdi et al. (2008)),
- les tractographies probabilistes
 - par ligne de courant (Perrin et al. (2005), Parker et Alexander (2005) par exemple),
 - ou bayésiennes (Behrens et al. (2007) par exemple)
- les tractographies globales (Poupon (1999), Jbabdi (2007) par exemple).

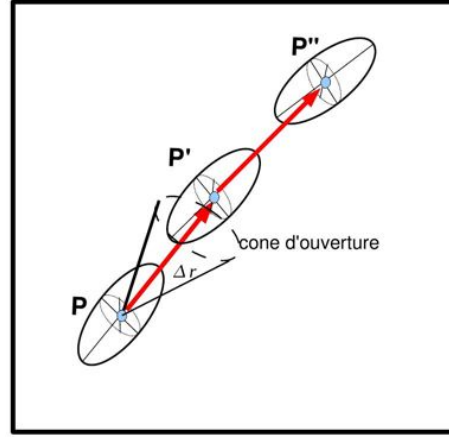
Parmi ces nombreuses méthodes, nous avons testé dans cette thèse deux approches partant sur le même principe des lignes de courant : soit en mode déterministe, soit en mode probabiliste. Nous allons présenter brièvement le principe et les avantages et inconvénients de ces méthodes maintenant.

Déterministe La tractographie déterministe par ligne de courant consiste à initialiser des graines d'où les tracts seront reconstruits. A partir de chaque graine prend naissance un tract qui sera prolongé en suivant pas à pas la direction la plus probable du modèle local.

Cette démarche a le mérite d'être simple et rapide mais elle est très sensible au bruit, et ne gère pas les croisements de fibres. Il existe plusieurs alternatives possibles dont la tractographie probabiliste par ligne de courant que nous avons utilisée dans notre thèse.

Probabiliste La tractographie probabiliste par lignes de courant reconstruit aussi pas à pas les tracts à partir de graines initiales. A la différence de la tractographie déterministe, à chaque pas, on suit une direction prise aléatoirement dans un cône d'ouverture (cf. Fig. 2.11) selon le modèle de diffusion local. La méthode de tractographie utilisée dans cette thèse a été

FIG. 2.11 – Principe de la tractographie par ligne de courant avec l'exemple du tenseur de diffusion. A partir d'une graine P , le tract est reconstruit pas à pas. Dans le cas déterministe, à chaque pas (Δr), la direction principale du modèle local est suivie. Dans le cas probabiliste, une direction est prise aléatoirement dans le cône d'ouverture. (source : Poupon, cours 2010)



proposée par Perrin et al. (2008) et consiste en une approche probabiliste de type Monte-Carlo avec une régularisation de la courbure.

Cette méthode nécessite en entrée un masque de propagation. Chaque voxel de ce masque est sur-échantillonné en un nombre n de points de départ des particules ou graines initiales. Chaque particule a une vitesse initiale suivant la direction principale du modèle local, puis cette particule se déplace dans le champ d'ODFs selon le procédé suivant.

En notant $\mathbf{p}(t)$ et $\mathbf{v}(t)$ la position et la vitesse de la particule au temps t , la particule se déplacera au temps $t + \delta t$ en $\mathbf{p}(t + \delta t)$ avec :

$$\mathbf{p}(t + \delta t) = \mathbf{p}(t) + \mathbf{v}(t) \cdot \delta t \quad (2.9)$$

A chaque pas, la vitesse sera recalculée selon l'équation suivante :

$$\mathbf{v}(t + \delta t) = \alpha \mathbf{v}_q + (1 - \alpha) \mathbf{v}(t) \quad (2.10)$$

Cette nouvelle vitesse combine linéairement un terme d'inertie (la vitesse du pas précédent), avec une nouvelle vitesse \mathbf{v}_q calculée en fonction d'une force provenant du modèle local de diffusion.

L'orientation \mathbf{v}_q est choisie aléatoirement dans un demi-cône défini à partir de la direction incidente $\mathbf{v}(i)$ suivant une distribution de probabilité $F_q(\mathbf{v}(t))$. Cette distribution de probabilité $F_q(\mathbf{v}(t))$ correspond à l'ODF (d-ODF ou f-ODF) modélisée restreinte sur ce demi cône. Notons qu'une version déterministe de cette régularisation est possible, si l'on prend la direction principale du modèle local pour \mathbf{v}_q (Perrin et al. 2005).

Le **paramètre** α correspond à l'écart type de la distribution de l'ODF, normalisé par son maximum dans le champ d'ODFs et représente une mesure de l'**anisotropie**. En effet, lorsque la diffusion est isotrope, l'écart type considéré est faible et le paramètre α l'est donc aussi. Dans ce cas, l'algorithme favorise la direction incidente. Lorsque la diffusion est anisotrope, α est grand et l'algorithme favorise les données de diffusion. L'algorithme est expliqué en détails dans la thèse de Muriel Perrin (Perrin 2006).

Cette approche de type tractographie probabiliste par ligne de courant a l'avantage d'être robuste au bruit et au volume partiel (permettant ainsi de traverser un croisement de fibres). Elle est en revanche coûteuse en calcul, mais est parallélisable donc elle reste efficace.

2.4 RÉCAPITULATIF DES DONNÉES PRÉTRAITÉES DISPONIBLES

Dans cette thèse, nous avons traité les données de la base NMR décrites précédemment au début de ce chapitre (cf. partie 2.1.2). Nous allons présenter dans cette partie un récapitulatif des différents types de données qui ont été générés avec les outils présentés dans la partie précédente.

Nous avons traité ces données selon deux chaînes de traitement différentes en commençant par une analyse intra-sujet des données (Roca et al. 2009a) puis en poursuivant par une analyse inter-sujets (Roca et al. 2010a).

2.4.1 Premiers traitements

Les traitements présentés ici ont servi pour deux publications dans des conférences (Roca et al. 2009a;b).

Données anatomiques Dans cette partie, nous avons précisé le contexte. Les données anatomiques ont été traitées à l'aide de Brainvisa selon la chaîne de traitement décrite précédemment aboutissant à l'extraction des maillages de l'interface Gris/Blanc des sujets.

Données de diffusion Nous avons utilisé les méthodes disponibles au laboratoire, pour bénéficier du conseil et de l'expertise locale.

Modèle local Un champ de Q-ball analytique est calculé à partir des données de diffusion. L'ODF de diffusion est ici reconstruite à partir d'une décomposition en harmoniques sphériques du signal (Descoteaux et al. 2007).

Les tracts Les tracts sont reconstruits par l'algorithme de tractographie déterministe par ligne de courant de Ptk qui inclut une régularisation de la courbure Perrin et al. (2005). La tractographie est initiée avec deux particules par voxels, dans un masque de la matière blanche à la résolution de l'image anatomique qui tient compte des sillons corticaux (Perrin et al. 2008). La tractographie est stoppée quand les particules sortent du masque de propagation, quand l'angle entre les deux derniers mouvements dépasse 30° , ou quand la longueur du tract dépasse 200 mm.

Les tract de longueur inférieure à 20 mm sont éliminés, donnant alors un nombre total moyen de 1,5 millions de tracts par sujet.

2.4.2 Seconds prétraitements

Les traitements présentés ici ont servi pour deux publications dans des conférences (Roca et al. 2010a;b). La figure 2.12 récapitule les différents prétraitements effectués ainsi que les données d'entrée et les données générées.

Données anatomiques Les données anatomiques ont été traitées avec Freesurfer par Alan Tucholka selon les traitements présentés précédemment (cf. partie 2.3.2.2). Ils aboutissent alors à des maillages de l'interface

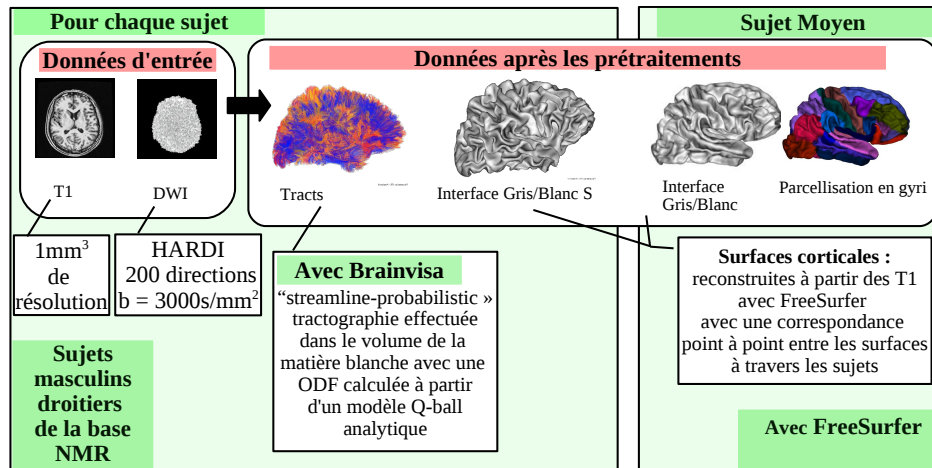


FIG. 2.12 – Données prétraitées disponibles servant à construire la matrice de connectivité avec : - pour chaque sujet les tracts, l'interface gris/blanc ; - pour l'ensemble des sujets : un maillage moyen de l'interface gris/blanc et la parcellisation en gyri de ce maillage.

Gris/Blanc qui présentent une correspondance point à point à travers les sujets. Une parcellisation en gyri est aussi disponible, ainsi qu'une surface corticale moyenne qui servira pour la visualisation.

Données de diffusion

Modèle local Les images de diffusion ont été traitées à l'aide de Ptk. Elles sont préalablement débruitées à l'aide d'un filtre adapté au bruit Rician (Descoteaux 2008). La fonction de distribution des orientations des fibres est ensuite reconstruite dans chaque voxel à partir des données Q-ball (SH-Qball Descoteaux et al. (2007)) à l'aide d'une déconvolution sphérique (Descoteaux et al. 2009).

Tracts Les fibres sont ensuite reconstruites à partir de l'algorithme de tractographie probabiliste par ligne de courant avec régularisation de la courbure décrit précédemment (Perrin et al. 2008). Le masque de propagation ainsi que les paramètres de tractographie sont les mêmes que pour le cas précédent à la différence du nombre de particules par voxel qui est de 10 et l'angle de courbure maximum qui est fixé à 45°. Les tracts sont de plus stoppés quand ils dépassent une longueur de 500 mm et non plus 200 mm. En éliminant les tracts inférieurs à 5 mm nous obtenons ainsi un total de 7 millions de tracts par sujet environ.

2.5 CONCLUSION

Dans cette thèse, nous allons étudier le connectome dans un cadre surfacique et structurel, i.e. en prenant comme support la surface corticale et comme connexions les tracts, reconstructions des faisceaux de neurones. La construction de ce connectome nécessite deux modalités d'acquisition : l'IRM anatomique à partir de laquelle sera reconstruite la surface corticale, et l'IRM de diffusion qui permettra de reconstruire une approximation des faisceaux, les tracts.

Dans ce chapitre, nous avons rappelé les principes de l'imagerie par Résonance Magnétique pour les deux modalités précédentes. Pour chacune de ces modalités, nous avons ensuite présenté les principaux prétraitements qui ont été effectués sur les données traitées pendant ma thèse.

Dans le chapitre suivant (Chapitre 3), premier chapitre faisant partie des contributions de cette thèse, ces deux modalités seront combinées au travers de nouveaux outils pour construire un connectome structural surfacique et le visualiser.

Dans la suite de cette partie, à partir de ces outils, nous présenterons dans les autres chapitres des pistes pour analyser ce connectome en proposant : une manière de le simuler (Chapitre 4), des façons de réduire et simplifier les données de connectivité sous-jacentes (Chapitre 5), une méthode pour parcelliser le cortex d'un sujet entièrement à partir de ces données (Chapitre 6), ainsi qu'une méthode parcellisant une région du cortex pour un groupe de sujets (Chapitre 7).

Deuxième partie

Contributions

OUTILS SPÉCIFIQUES AU CONNECTOME ANATOMIQUE SURFACIQUE CORTICAL

Pour étudier le **connectome structurel et surfacique**, il est nécessaire d'avoir les outils pour le **construire**, le **traiter** et le **visualiser**. Dans ce contexte, deux modalités d'imagerie sont combinées : l'imagerie anatomique et l'imagerie de diffusion, comme nous venons de le voir au dernier chapitre de la partie précédente. De plus, l'étude de ce type de connectome en particulier est assez récente et il n'existait ainsi pas d'outils logiciels validés et standard lors du début de ma thèse pour effectuer une telle étude. Des développements logiciels ont donc été nécessaires pendant cette thèse pour réaliser de nouvelles fonctionnalités spécifiques à l'étude du connectome structurel surfacique.

Dans ce chapitre, nous allons ainsi décrire trois exemples de telles fonctionnalités, qui seront bientôt intégrées dans Brainvisa. Un exemple d'utilisation étant une chaîne de traitements allant de la création des noeuds du connectome par sur-segmentation d'une parcellisation en gyri, à la construction de la matrice de connectivité correspondante de la même façon que Hagmann et al. (2008). Des outils originaux de visualisation de cette matrice ont de plus été implémentés, permettant de mieux appréhender la répartition spatiale des connexions sur l'intégralité du cortex.

SOMMAIRE

4.1	INTRODUCTION	77
4.2	CONSTRUCTION DES DONNÉES SIMULÉES	78
4.2.1	Construction du sujet <i>modèle</i>	80
4.2.1.1	Les parcelles et leur support	80
4.2.1.2	Les connexions des parcelles : Graphe $G_{\mathcal{A},bin}^0$	80
4.2.1.3	Les connexions entre les points du maillage	81
4.2.2	Construction d'une base de sujets	84
4.3	DONNÉES UTILISÉES POUR LA VALIDATION	86
4.3.1	Choix des paramètres	86
4.3.1.1	Création du sujet <i>modèle</i>	86
4.3.1.2	Création de différents jeux de données	87
4.3.2	Deux bases de données pour illustrer les méthodes	88
4.3.2.1	Cas optimal	88
4.3.2.2	Cas extrême	88
4.4	CONCLUSION	90

3.1 INTRODUCTION

Durant ma thèse, je me suis appuyée sur les bibliothèques python et C++ de traitement surfacique et des données de diffusion développées par mon laboratoire au sein du logiciel Brainvisa pour satisfaire mes objectifs en les complétant si besoin. Ce travail ne constitue pas en soi un travail de recherche, mais il a été nécessaire et utile pour la suite de ma thèse. Les fonctionnalités nouvelles développées seront de plus disponibles dans ce même logiciel.

En voici ci-dessous une brève liste avec des exemples d'utilisation dans le cadre de ma thèse. Nous les décrirons plus en détails dans la suite du chapitre.

1. Approximation de l'algorithme de Voronoï centroïdal surfacique :
 - application à la sur-segmentation des gyri du cortex, fournissant ainsi une parcellisation qui peut servir pour la construction du connectome (expliqué dans ce chapitre);
 - construction d'un ensemble de parcellisations simulées du cortex correspondant à une base de sujets avec contrôle de la variabilité inter-individuelle des parcelles (au niveau de leur localisation); (cf. Chapitre 4);
 - découpage automatique du cortex dans le cas de la parcellisation intra-sujet du cortex entier (cf. Chapitre 6);
2. Construction de la matrice de connectivité corticale surfacique : reprise du travail effectué par Cathier et Mangin (2006) pour la construction de la matrice de connectivité corticale en le complétant (nouveau mode avec le calcul de l'intersection des tracts avec la surface corticale);
3. Visualisations avec Anatomist (présentées dans ce chapitre) :
 - de la matrice de connectivité : nouveau type d'objet permettant notamment de mieux visualiser l'aspect spatial des connexions pour l'ensemble d'un hémisphère;
 - des tracts associés : nouveau mode de fusion et de sélection des noeuds et des relations d'un graphe.

A l'aide de ces outils, il est ainsi possible, à partir d'un maillage du cortex parcellisé en gyri et des résultats d'une tractographie, de construire et de visualiser le connectome structurel surfacique associé à une sur-segmentation de cette parcellisation en gyri.

VOCABULAIRE

Pour un sujet, les données d'entrée de ces traitements sont :

- la **surface corticale** représentée par \mathcal{V} et constituée de V points $\{v_1, \dots, v_V\}$;
- les **tracts** qui sont des approximations numériques des fibres réelles, reconstruits par tractographie.

Dans ce chapitre, nous allons notamment décrire comment construire des parcelles sur la surface corticale, ainsi que les profils de connectivité les caractérisant.

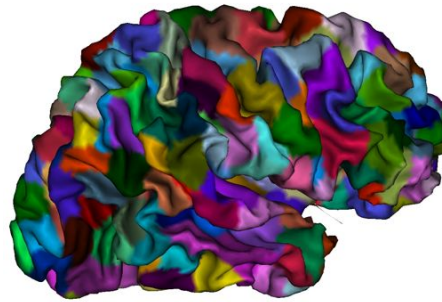
Une **parcelle** sera définie comme une zone connexe de points (Fig.3.1(a)). Dans les chapitres 6 et 7 l'objectif sera d'identifier des parcelles ayant des profils de connectivité homogènes.

Le **profil de connectivité** d'un point sera défini par un vecteur représentant sa force de connexion vers l'ensemble des autres points du cortex et pourra se représenter sous la forme d'un maillage texturé (Fig.3.1(b)). La **matrice de connectivité** corticale complète résultera de la concaténation des profils de connectivité de l'ensemble des points de la surface corticale.

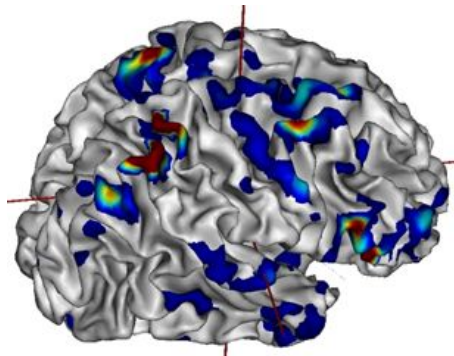
Si l'on considère une parcellisation du cortex, le profil de connectivité d'une parcelle pourra alors prendre deux formes :

- soit le profil complet : ce sera un profil de taille V représentant sa force de connexion vers chacun des points du maillage,
- soit le profil vers les autres parcelles : représentant sa force de connexion vers l'ensemble des autres parcelles.

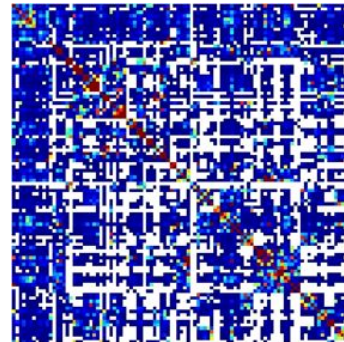
La concaténation des profils de chaque parcelle (vers les autres parcelles) constitue alors une **matrice de connectivité régionale**, i.e représentant les connexions entre les parcelles (Fig.3.1(c)).



(a) Parcelles



(b) Profil



(c) Matrice

FIG. 3.1 – Les données dont la construction sera détaillée dans ce chapitre : (a) Segmentation en parcelles par Voronoï Centroidal (cf. 3.2.1); (b) Profil de connectivité d'une parcelle; (c) Exemple de matrice de connectivité entre les parcelles.

3.2 PARCELLISATIONS ALÉATOIRES DU CORTEX

Pour l'étude et la construction d'un connectome surfacique, il est parfois utile d'avoir un outil de segmentation automatique de la surface corticale en un nombre variable de régions : (i) pour construire le connectome à différentes échelles avec une correspondance entre les noeuds des sujets (Hagmann et al. 2008), (ii) pour réduire la dimension des données de connectivité (cf. partie sur l'étude des profils de connectivité 5.3.2), ou (iii) pour découper la surface corticale en vue d'un clustering itératif du cortex (cf. partie Parcellisation intra-sujet du cortex 6 et Roca et al. (2009b)).

Des outils fournissant une parcellisation automatique du cortex en fonction de l'anatomie sulco-gyrale sont disponibles (dans Brainvisa et Freesurfer notamment, cf. partie Définition des noeuds du cortex 1.2.1.2). Ils aboutissent à une parcellisation en gyri du cortex.

Dans le cadre de cette thèse, il a été nécessaire de segmenter automatiquement de manière plus fine le cortex avec comme contraintes le fait que les **parcelles** forment une **partition de la région d'intérêt** initiale à parcelliser, et qu'elles soient **uniformément réparties** sur la surface (et donc de tailles similaires).

Ceci peut être effectué :

- soit de façon complètement **aléatoire** sans tenir compte de la géométrie de la surface corticale,
- soit en tenant compte d'une segmentation préalable : **Sur-segmentation d'une parcellisation initiale** (en gyri par exemple (Hagmann et al. 2008, Gigandet 2009)).

Nous avons implémenté ces types de parcellisations surfaciques qui seront bientôt disponibles dans Brainvisa. Celles-ci n'ont pas besoin d'un système de coordonnées surfaciques, elles tiennent seulement compte de la distance géodésique entre les points constituant la surface corticale ou de régions d'intérêt sur cette surface.

3.2.1 Diagramme de Voronoï centroïdal surfacique

Pour ce faire, nous avons utilisé un algorithme de diagramme de Voronoï surfacique géodésique (Cachia et al. 2003). Pour obtenir des parcelles uniformément réparties sur la surface corticale, une procédure itérative est mise en place (algorithme de Voronoï centroïdal). La description de ces algorithmes est très bien décrite dans la thèse de Flandin (2004). Nous allons en décrire brièvement les principes sous-jacents.

Définition 3.1 Soit un ensemble fini $\mathcal{V} = \{v_i\}_{i=1}^V$, K éléments de \mathcal{V} appelés *graines* $\{c_k\}_{k=1}^K$ et une mesure de distance d sur \mathcal{V} . On appelle **parcellisation de Voronoï** de \mathcal{V} en K **cellules de Voronoï** la partition de \mathcal{V} en K sous-ensembles $\{\mathcal{V}_k\}_{k=1}^K$ définis par :

$$\mathcal{V}_k = \{x \in \mathcal{V} \mid d(x, c_k) < d(x, c_l), \forall k \neq l\} \quad (3.1)$$

Définition 3.2 Soit un ensemble fini $\mathcal{V} = \{v_i\}_{i=1}^V$, K éléments de \mathcal{V} appelés *graines* $\{c_k\}_{k=1}^K$ et une mesure de distance d sur \mathcal{V} . Une partition $\{\mathcal{V}_k\}_{k=1}^K$ de \mathcal{V} constitue une **parcellisation de Voronoï centroïdale** si :

- les $\{\mathcal{V}_k\}_{k=1}^K$ sont des cellules de Voronoï étant donné les graines $\{c_k\}_{k=1}^K$,
- les graines $\{c_k\}_{k=1}^K$ sont les barycentres des régions $\{\mathcal{V}_k\}_{k=1}^K$.

La construction de la parcellisation de Voronoï centroïdale d'un domaine est effectuée par l'algorithme des k-moyennes ou **K-means** (Hartigan et Wong 1979) :

Algorithme 1: Algorithme K-means

Initialisation des K centroïdes $\{c_k\}_{k=1}^K$ des K classes

Itération jusqu'à convergence :

Classification des $\{v_i\}_{i=1}^V$ en fonction du $\{c_k\}$ le plus proche

Mise à jour des centroïdes $\{c_k\}_{k=1}^K$

L'étape de mise à jour des centroïdes nécessite le calcul du barycentre de chaque classe. Dans le cas d'une distance euclidienne, une formulation explicite existe. Lorsque la distance utilisée est la distance géodésique le long de la surface corticale, il n'y a pas de formulation explicite. Pour simplifier dans notre cas nous avons pris comme approximation du barycentre le point v_i de la classe dont la distance aux frontières de la classe est maximale. C'est une approximation car il n'est pas forcément unique mais dans notre cas, nous avons testé et au bout d'un nombre raisonnable d'itérations (5) les classes obtenues sont satisfaisantes (approximativement uniformément réparties à la surface). Lors de cette démarche, pour chaque classe une seule carte de distance à partir des points frontières est calculée : le calcul de la distance géodésique entre les points de la surface n'est pas requis ce qui allège les calculs.

3.2.2 Applications

Durant cette thèse nous avons utilisé ce type de parcellisation dans d'autres cas de figure, comme par exemple :

- Segmentation du cortex avec contraintes sulco-gyales (cf. développement ci-dessous 3.2.2) ;
- Construction d'une base simulée de parcellisations du cortex avec contrôle de la variabilité inter-individuelle (cf. Chapitre 4),
- Segmentation automatique du cortex dans le cas d'une parcellisation itérative intra-sujet du cortex entier (cf. Chapitre 6).

Segmentation du cortex avec contraintes sulco-gyales En suivant la même démarche que dans l'article d'Hagmann et al. (2008) (cf. partie 1.1.4.2), j'ai créé une parcellisation de la surface corticale en 200 régions, en prenant comme contrainte la parcellisation en régions anatomiques réalisée par Freesurfer (cf. Desikan et al. (2006) et partie 2.3.2.3).

Chaque région anatomique de cette parcellisation est segmentée elle-même en cellules de Voronoï approximativement centroïdales suivant le principe précédemment décrit. Pour chaque région A_i de la parcellisation initiale, un nombre K_i de graines de Voronoï est choisi en fonction de l'aire de cette région $surf_{A_i}$ et d'un paramètre fixé $surf_{min}$ d'unité le cm^2 :

$$K_i = \text{Partie entière de } \left(\frac{surf_{A_i}}{surf_{min}} \right)$$

En prenant $surf_{min} = 6cm^2$, nous obtenons ainsi la parcellisation de la figure 3.2.

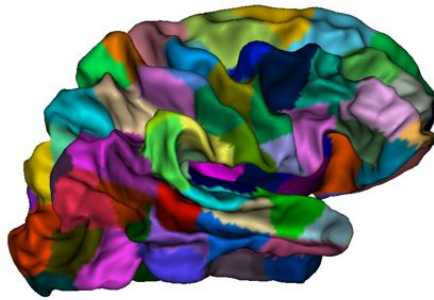


FIG. 3.2 – Diagramme de Voronoï centroïdal. Exemple de parcellisation en 200 régions de la surface corticale de l'hémisphère droit. Le maillage de cette surface utilisé ici est le maillage moyen de la base de sujets de cette thèse, fourni par Freesurfer (cf. 2.3.2.2).

Ce genre de parcellisation du cortex a servi ainsi à effectuer une réduction de dimension des profils de connectivité corticaux avant une analyse en composantes principales (cf. partie 5.3).

De plus, cela peut fournir une manière de construire les noeuds du connectome. Lorsque les maillages de la surface corticale pris sont ceux de Freesurfer avec une correspondance entre les sujets, il est alors possible de moyenner la matrice de connectivité associée (cf. Hagmann et al. (2008)). Nous allons voir dans la section suivante comment construire une telle matrice de connectivité à partir des parcelles ainsi obtenues. Cette matrice peut ensuite être moyennée à travers un groupe de sujets.

Cette démarche n'apporte rien de nouveau par rapport à l'approche d'Hagmann, si ce n'est que le moyennage est effectué sur plus de sujets (le double). Nous proposons de plus une visualisation nouvelle et originale du connectome ainsi construit (cf. partie Visualisation 3.4).

3.3 MATRICE DE CONNECTIVITÉ CORTICALE

Pour chaque sujet, à partir de la surface du cortex et des tracts issus d'une tractographie sur le cerveau entier, on peut construire la *matrice de connectivité corticale*. Il faut bien noter qu'il n'y a pas une définition de matrice mais une multiplicité de définitions, dont nous allons faire une liste (non exhaustive) dans un souci de clarification. Nous expliquerons ensuite comment les construire.

3.3.1 Différentes définitions

Notons $S = (\mathcal{V}, \mathcal{E}_{\text{spatial}})$ la surface du cortex constituée d'un ensemble de V points ou points $\mathcal{V} = \{v_1, \dots, v_V\}$ et d'un ensemble de relations spatiales les reliant $\mathcal{E}_{\text{spatial}}$. Soit $C_V \in M_{(\mathcal{V}, \mathcal{V})}(\mathbb{R})$ la matrice de connectivité corticale du cortex $S = (\mathcal{V}, \mathcal{E}_{\text{spatial}})$. $C_V(i, j)$ représente la force de connexion reliant les points v_i et v_j de \mathcal{V} , et peut être soit un nombre de tracts, soit une densité de connexion (nombre de tracts par unité de surface), soit un nombre pondérée en fonction de paramètres spécifiques aux tracts (longueur, anisotropie moyenne le long du tract, etc.).

On peut distinguer différents types de matrices de connectivité corticales :

- la **matrice des connexions** ou matrice de connectivité binaire : $C_{V, \text{bin}} \in M_{(\mathcal{V}, \mathcal{V})}(\{0, 1\})$ où chaque connexion entre deux entrées de la matrice est symbolisée par un 1, en définissant l'existence d'une

- connexion par l'existence d'un nombre minimal de tracts connectant les deux régions concernées, cette matrice étant donc paramétrée par une valeur de seuil (notée *NombreDeTractsMinimum* par la suite);
- **la matrice des tractogrammes** ou matrice pondérée à valeurs entières, appelée dans cette thèse **matrice de connectivité** : $C_V \in M_{(V,V)}(\mathbb{N})$ où chaque connexion est représentée par le nombre de tracts associés;
 - **la matrice de force pondérée des connexions** ou matrice pondérée à valeurs réelles : $C_{V,*} \in M_{(V,V)}(\mathbb{R})$ où chaque connexion est représentée par un poids calculé en fonction de différents paramètres issus des tracts (nombre de tracts, longueur moyenne associée, probabilité des tracts par exemple); un lissage géodésique gaussien peut aussi être appliqué à la matrice;
 - **la matrice des densités de connexion** : $C_{V,density}$ où chaque connexion est représentée par un nombre de tracts par unité de surface, cf. partie 4.2 décrivant les données simulées pour une définition plus détaillée).

3.3.2 Construction de la matrice

A partir des extrémités des tracts Pour construire la matrice de connectivité C_V , une première approche (Cathier et Mangin 2006) consiste à regarder les extrémités de chaque tract. Ainsi, pour chaque tract, les points v_i et v_j appartenant au maillage de la surface corticale (V) les plus proches respectivement des deux extrémités E_i et E_j du tract sont identifiés. Les entrées $C_V(i, j)$ et $C_V(j, i)$ sont alors incrémentées de 1 en conséquence (un tract en plus relie P_i et P_j , cf. Fig.3.3).

A partir des points d'intersection des tracts avec V Une autre approche consiste à calculer pour chaque tract ses deux points d'intersection P_i et P_j avec le cortex. En pratique, les points d'intersection P_i et P_j d'un tract avec le maillage de la surface corticale (s'ils existent) coïncident rarement avec deux points v_i et v_j de la surface. La plupart du temps, le point d'intersection P_i va être dans un polygone du maillage. Nous approximerons alors P_i par le point v_k le plus proche de P_i faisant partie des sommets du polygone où se trouve P_i . On peut alors incrémenter $C_V(i, j)$ et $C_V(j, i)$ de 1.

Il est ensuite possible de seuiller cette matrice et de la binariser pour obtenir $C_{V,bin}$. Nous obtiendrons alors :

$$\forall (i, j) \in \{1, M\}^2, C_V(i, j) = \begin{cases} 1 & \text{si } C_{V,bin}(i, j) > \text{NombreDeTractsMinimum} \\ 0 & \text{sinon.} \end{cases} \quad (3.2)$$

3.3.3 Connectome entre parcelles et pour un groupe de sujets

Matrice de connectivité entre parcelles La matrice de connectivité corticale C_V décrite précédemment comporte le nombre de tracts existant

entre les points v_i du maillage de la surface corticale. Lorsque l'on dispose d'une parcellisation $\mathcal{A} = \cup_{p=1..P} A_p$ de cette surface en parcelles, il est alors possible de construire à partir de C_V la matrice de connectivité $C_{\mathcal{A}} \in M_{(P,P)}(\mathbb{N})$ où $C_{\mathcal{A}}(p, q) =$ nombre de tracts reliant les parcelles A_p et A_q . $C_{\mathcal{A}}$ est construite à partir de C_V en sommant pour chaque parcelle A_p les contributions des points v_i la constituant.

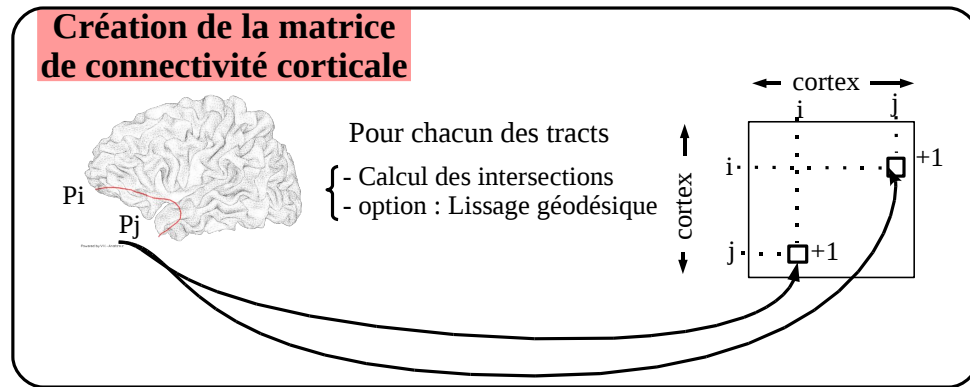


FIG. 3.3 – Calcul de la matrice de connectivité corticale. Pour chaque sujet, les données d'entrée sont les tracts et le maillage de l'interface gris/blanc (cf.2.12).

Matrice de connectivité d'un groupe de sujets Pour un groupe de sujets, nous avons vu précédemment qu'il est possible d'obtenir des maillages de l'interface Gris/Blanc ayant une correspondance point à point entre les sujets ainsi qu'un maillage moyen du groupe et une parcellisation en gyri, à l'aide du logiciel Freesurfer (cf. 2.3.2.2). En effectuant une sur-segmentation de ces gyri à partir d'une parcellisation de Voronoï centroïdale surfacique (cf. Section 3.2.2), il est alors possible d'obtenir une sur-segmentation du cortex avec contraintes sulco-gyrales. Les parcelles se correspondent d'un sujet à l'autre, la qualité de cette correspondance dépendant bien sûr de l'étape de recalage effectuée lors de la construction des maillages par Freesurfer.

A partir de ces parcellisations, la matrice de connectivité entre parcelles est construite pour chaque sujet conformément au paragraphe précédent. Une matrice de connectivité au niveau du groupe de sujets peut ensuite être calculée en effectuant la moyenne de ces matrices. Nous verrons dans la section suivante comment améliorer la visualisation d'une telle matrice.

3.4 VISUALISATION DE LA MATRICE DE CONNECTIVITÉ

3.4.1 Outils existants

Il existe différents types d'outils de visualisation des matrices de connectivité, parmi lesquels nous pouvons citer (liste non exhaustive) :

- logiciels de visualisation 3D de tracts : Anatomist, trackvis, MedInria ;
- logiciels de visualisation 3D de graphes de connectivité : connec-

tome Viewer (cf. Fig. 3.4), utilisant Mayavi et le module networkx de python ;

- images 2D (exemples dans Hagmann et al. (2008)) :
 - sous forme de matrice classique (cf. Fig 3.5.A) ;
 - “Kamada-Kawai force-spring layout” (cf. Fig 3.5.B) ;
 - graphe squelette (cf. Fig 3.5.C).

FIG. 3.4 – Représentation d’une matrice de connectivité avec ConnectomeViewer. Chaque point représente une région du cortex. Les segments reliant les points représentent les connexions. Des couleurs peuvent leur être attribuées (en fonction de la force de connexion existant entre les différentes régions par exemple).

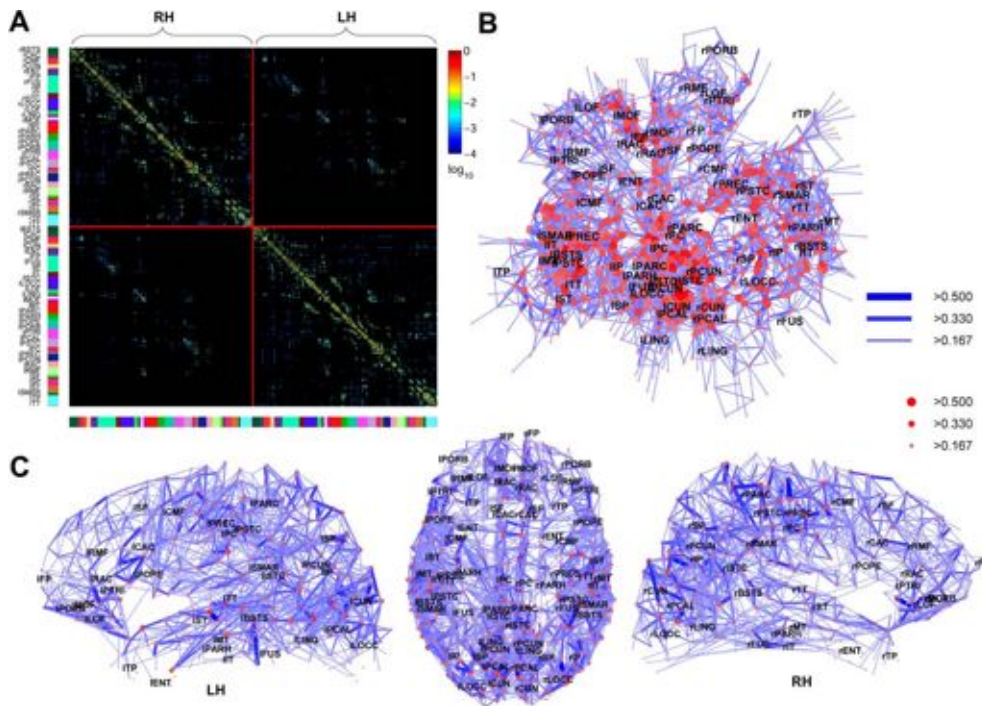
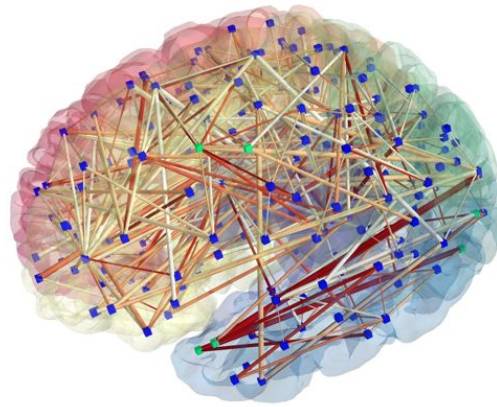


FIG. 3.5 – Différents types de visualisation 2D du connectome : (A) Matrice de connectivité entre les paires de $n = 998$ régions, (B) Représentation de Kamada-Kawai du squelette de la connectivité, (C) Vues dorsale et latérale du squelette de la connectivité. (source : Hagmann et al. (2008)).

3.4.2 Contributions personnelles

Représentation surfacique interactive de la matrice de connectivité

Pour mieux visualiser l'aspect spatial des connexions, en combinant différentes échelles, j'ai mis au point avec Denis Rivière un autre moyen de visualisation. Considérons la surface corticale S composée de N points. Considérons une parcellisation du cortex en I régions et deux matrices de connectivité associées :

- la *matrice régionale régions-régions*, de taille (I, I) , graphe de connexions entre les régions, une ligne i représentant le profil de connectivité $p_{A_i \rightarrow (A_j)_{j=1..I}}$ de la région A_i vers l'ensemble des régions $(A_j)_{j=1..I}$;
- la *matrice mixte régions-points*, de taille (I, N) , chaque ligne étant le profil de connectivité $p_{A_i \rightarrow (v_j)_{j=1..N}}$ d'une région vers les N points (ou *points*) du cortex.

Au lieu de représenter chaque région par un point dans l'espace 3D comme précédemment (Fig. 3.5), chaque région ou parcelle est représentée par la partie de la surface corticale associée dans l'espace 3D, avec un gonflement global du cortex pour simplifier la visualisation (gênée par les sillons et gyri). Pour chacune de ces parcelles, un cerveau miniature représentant son profil de connectivité est ajouté au dessus de cette parcelle. Ce profil peut être celui de la *matrice régionale* $p_{A_i \rightarrow (A_j)_{j=1..I}}$ (Fig.3.6), ou celui de la *matrice mixte régions-points* $p_{A_i \rightarrow (v_j)_{j=1..N}}$. Ce mode de visualisation 3D est **interactif** : il est possible de tourner la surface corticale dans l'espace 3D pour visualiser l'ensemble du cortex.

Combinaison de différentes échelles Cette visualisation a l'avantage de combiner différentes échelles : la visualisation des profils $p_{A_i \rightarrow (v_j)_{j=1..N}}$ de chaque région A_i permet de mieux comprendre la répartition à l'échelle des *points* des connexions au sein des parcelles A_i , et ceci pour chacune de ces régions.

La répartition spatiale des connexions est en outre plus facile à appréhender que dans les approches précédentes où l'on est vite perdu avec la multitude de segments représentant les connexions, que ce soit au niveau régional (graphe de connexion entre les régions) ou à deux échelles avec les *profils mixtes régions-points*. Sur la figure 3.6 qui représente une visualisation du type *régions-régions* d'un atlas macroscopique du connectome, il apparaît ainsi très nettement que les connexions courtes ont des poids plus importants que les connexions longue distance. Cette visualisation permet de se rendre compte de la prépondérance des tracts courts sur les tracts longs dans les données : autour de chaque zone, les zones les plus fortement connectées sont les zones proches.

Nouvelle représentation 3D en graphe interactif des tracts et des parcelles surfaciques associées à la matrice de connectivité

Avec l'aide de Denis Rivière, nous avons aussi créé un nouveau type de graphe dans Anatomist, donc les noeuds représentent des morceaux de

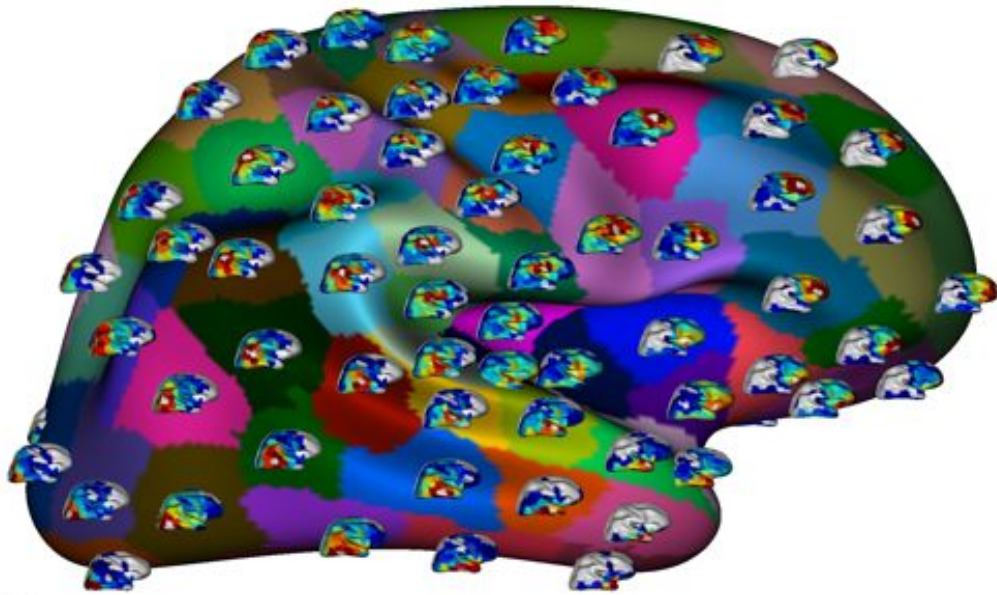


FIG. 3.6 – Exemple d’atlas macroscopique du connectome humain : Représentation de la matrice de connectivité d’une segmentation du cortex humain en 200 régions, moyennée à travers 10 sujets. Pour chaque région, un petit maillage du cortex est ajouté au-dessus du centroïde, représentant sa force de connexion vers les autres régions du cerveau (force croissante du bleu au rouge).

surface (sous forme de maillage) et les relations les tracts les reliant. Un mode de sélection de ce graphe a été alors implémenté, permettant entre autres :

- de sélectionner un morceau de surface et d’afficher les tracts en partant,
- de sélectionner plusieurs morceaux de surface et d’afficher les tracts les reliant.

Ce mode est très pratique pour visualiser les tracts associés à une matrice de connectivité correspondant à une parcellisation du cortex.

3.5 CONCLUSION

Dans ce chapitre a été décrite une chaîne complète de traitements construisant un connectome structurel surfacique régional à partir des surfaces corticales et des résultats d’une tractographie cerveau entier, avec comme étapes :

- Construction d’une parcellisation du cortex avec des contraintes sulco-gyrales ou non ;
- Calcul de la matrice de connectivité :
 - entre les points de la surface corticale,
 - entre les parcelles de la parcellisation précédente ;
- Visualisation originale du connectome structurel surfacique régional ainsi construit.

La visualisation finale est la partie la plus innovante de cette chaîne de traitements. En effet, les étapes précédentes ont déjà été développées par d’autres laboratoires (cf. Hagmann et al. (2008) avec une approche surfacique également). Il est cependant important de décrire toutes ces

étapes, étant donné que ces traitements ne se sont pas encore standardisés dans la communauté scientifique.

Chaque étape sera réutilisée dans la suite du manuscrit dans d'autres cas d'application originaux comme par exemple dans l'étude de la connectivité anatomique de la région englobant le SMA et le pré-SMA vers un ensemble de régions d'intérêt identifiées automatiquement avec :

- une visualisation des données de connectivité associées (cf. partie 5.2.5 et Fig. 5.20),
- un clustering de la région à partir de ces données en intra sujet, ou pour un groupe de sujets (cf. partie 7.2.4).

CONSTRUCTION DE CONNECTOMES SIMULÉS

Pour valider et comparer les études effectuées sur des données de connectivité réelles, comme la réduction de dimension des profils de connectivité et la parcellisation du cortex à partir de ces profils, construire des données de connectivité simulées est important. L'objectif de ce chapitre est de **construire des connectomes simulés pour un ensemble d'individus, en modélisant la variabilité inter-individuelle** de ces connectomes et **en contrôlant la densité des connexions**.

Pour ce faire nous utiliserons une représentation sphérique de la surface corticale, que nous parcelliserons en régions connexes ou parcelles regroupant un ensemble de points. A partir de ces données anatomiques une matrice de connectivité représentant le nombre de tracts reliant les parcelles sera construite. Une matrice de connectivité au niveau des points sera ensuite construite en tirant pour chaque tract connectant deux parcelles, les points de terminaison du tract dans chacune des parcelles parmi une distribution de probabilité que nous préciserons. Ainsi, les points de chacune de ces parcelles auront des profils de connectivité similaires. La densité des connexions contrôlera le degré de similarité entre ces profils. Pour contrôler la variabilité inter-individuelle, nous commencerons par construire le connectome d'un premier sujet appelé **sujet modèle** et nous fabriquerons à partir de ce sujet une **base de sujets dérivés** en dupliquant le sujet modèle tout en introduisant de la variabilité dans la position et la forme des parcelles.

Dans ce chapitre, nous décrirons dans une première partie les caractéristiques et la méthode de construction des données simulées. Dans une deuxième partie, nous présenterons les jeux de données dont nous nous servirons pour comparer les différentes méthodes de réduction des données de connectivité et les méthodes de parcellisation du cortex associées.

SOMMAIRE

5.1	INTRODUCTION	93
5.2	IDENTIFICATION AUTOMATIQUE DE RÉGIONS D'INTÉRÊT	95
5.2.1	Méthode Intra-sujet	95
5.2.2	Extension de la méthode à un groupe de sujets par moyennage	97
5.2.3	Données Simulées : cas optimal	98
5.2.4	Données Simulées : Base extrême	104
5.2.5	Essais sur les données réelles	106
	Région englobant le SMA et le pré-SMA	107
5.3	ANALYSE EN COMPOSANTES PRINCIPALES	111
5.3.1	Principe	111
5.3.2	Méthode	111
	1. Première réduction de dimension à partir d'une segmentation a priori	111
	2. Analyse en composantes principales de la matrice réduite \mathbf{X}'	113
5.3.3	Données Simulées : Cas optimal Template T_1	113
5.3.4	Données simulées : Cas extrême	117
5.3.5	Essais sur les données réelles	119
5.4	DIFFÉRENTES MESURES DE DISTANCE ENTRE PROFILS	122
5.4.1	Définitions	123
5.4.1.1	La corrélation	123
5.4.1.2	Distance euclidienne	123
5.4.1.3	Distance de Hausdorff	124
5.4.2	Données Simulées	125
5.4.2.1	Cas extrême	127
5.5	CONCLUSION	128

4.1 INTRODUCTION

Sporns et Zwi (2004) ont construit différents types de connectomes simulés dans le but de les comparer à des données réelles pour infirmer ou confirmer le caractère small-world des réseaux de connectivité cérébraux construits à partir des données de connectivité anatomique. Valdés-Sosa et al. (2005) utilisera le même genre de modèle pour simuler la connectivité fonctionnelle. Dans son étude, Sporns simule la surface corticale par un tore constitué de 40 points. Il fabrique ensuite trois types de connectomes : (i) un connectome présentant uniquement des connexions locales, (ii) un connectome présentant des connexions locales ajoutées à d'autres connexions (avec une distribution spatiale uniforme), (iii) un connectome constitué de deux régions ayant chacune des connexions locales ainsi que des connexions longue-distance (avec une distribution spatiale gaussienne dans les deux cas).

Dans ce chapitre, nous présentons une extension du dernier type de connectome en élargissant le nombre de régions (30 régions) que nous appellerons parcelles. Le but sera de simuler une **parcellisation du cortex ainsi que les données de connectivité reliant les parcelles**. Ces données simulées nous permettront d'évaluer des méthodes de parcellisation du cortex. En effet, à partir des données de connectivité simulées, nous chercherons à identifier les parcelles sous-jacentes et pourront alors évaluer le résultat obtenu à partir de la parcellisation simulée initialement. Dans un premier temps, le connectome simulé ne contiendra que des **connexions longues distances entre parcelles** et pas de connexions intra-parcelles.

Crippa et al. (2010) propose dans une étude récente une simulation des données de connectivité d'une région du cortex, dans le contexte d'une parcellisation de la région englobant l'aire motrice supplémentaire et le pré-SMA. Il construit ainsi 6 jeux de données, chacun simulant les données de connectivité d'une région composée de 250 points. Les 6 jeux de données sont construits de façon à avoir différents cas d'organisation des points, comme par exemple : 2 compartiments complètement séparés, 2 compartiments se recouvrant partiellement, un continuum de petits compartiments se recouvrant partiellement. L'objectif est de montrer que la méthode proposée peut identifier une transition graduelle entre les profils de connectivité des différents points.

Ici, notre objectif est différent. Contrairement à l'étude de Crippa qui est intra-sujet, notre but sera de comparer des méthodes de parcellisation inter-sujets. Nous ne simulerons donc non pas un connectome mais **plusieurs connectomes constituant ainsi une base de sujets** avec un **contrôle sur la variabilité inter-individuelle** dans la taille et la position **des parcelles**. Nous construirons ainsi différentes bases de sujets avec des variabilités inter-individuelles différentes. Nous ferons aussi **varier la densité des connexions** reliant les parcelles. Ceci nous permettra de comparer différentes approches de parcellisation de la surface corticale fonctionnant soit sur chaque sujet indépendamment, soit sur un groupe de sujets simultanément (cf. partie 7.2).

4.2 CONSTRUCTION DES DONNÉES SIMULÉES

La construction des données simulées est composée de deux phases :

1. Une première fabriquant un premier exemplaire : le “connectome” d’un sujet, appelé **sujet modèle** ;
2. Une deuxième construisant une **base de sujets dérivés** à partir de ce premier sujet, en introduisant une **variabilité inter-individuelle**.

Pour chaque sujet, la surface du cortex simulée est divisée en parcelles (regroupement de points de cette surface). Les connexions simulées entre ces parcelles forment un premier graphe, G_A à partir duquel est construit un graphe de connexions au niveau des points de la surface G_V (cf. Fig.4.1).

Une manière de comparer et de valider les techniques de parcellisation du cortex, sera d’essayer de retrouver les parcelles à partir du graphe de connexions au niveau des points du cortex G_V . Ceci sera effectué tout en contrôlant la variabilité inter-sujets et le bruit des données à partir d’un ensemble de paramètres (cf. Fig 4.2 et 4.3).

Nous allons maintenant présenter comment s’est effectuée la construction de ces données simulées en commençant par la première phase de construction du sujet modèle. Nous expliquerons ensuite comment des bases de sujets dérivés de ce sujet modèle peuvent être construites en injectant une variabilité inter-individuelle dans les parcellisations.

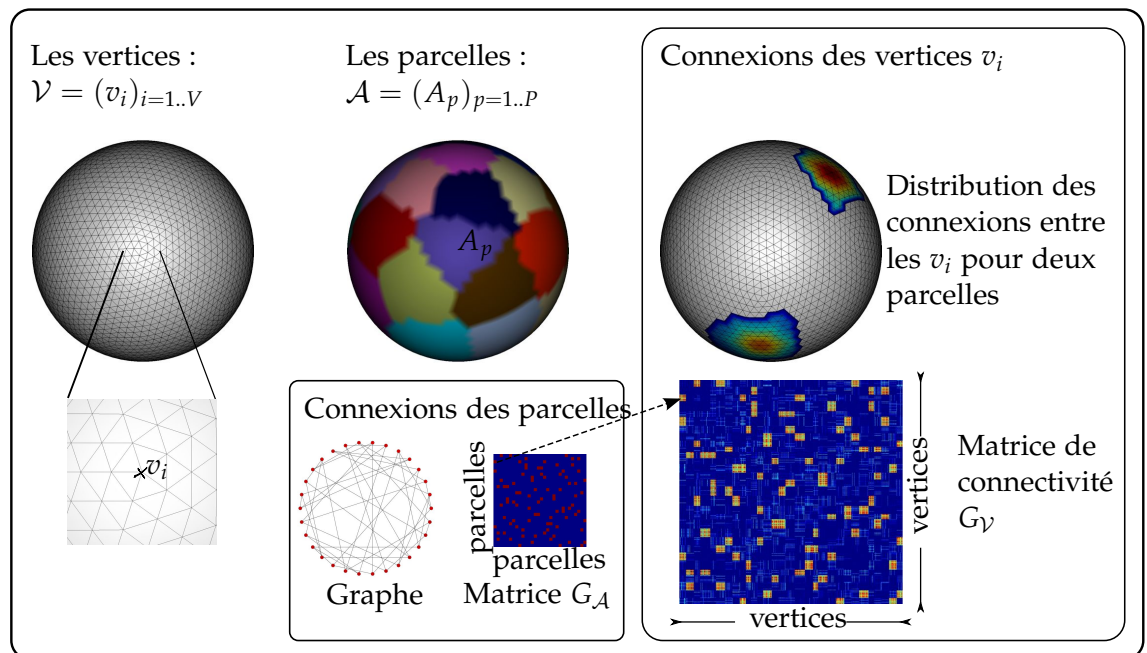


FIG. 4.1 – Données simulées pour un sujet : Le cortex est simulé par un maillage \mathcal{V} , constitué de points v_i regroupés en parcelles A_p . Deux types de matrices sont simulées : C_A et C_V , les matrices de connectivité entre les parcelles et entre les points du maillage du cortex entier respectivement.

Symboles	Noms des objets	Paramètres
$S = (\mathcal{V}, \mathcal{E}_{spatial})$ $\mathcal{V} = \{v_i\}_{i=1..V}$ $\mathcal{E}_{spatial}$	Surface du cortex points ou noeuds du maillage relations spatiales entre les noeuds du maillage (polygones)	aire $surf_S$ en mm^2 V : nombre de noeuds du maillage type et aire des polygones \rightarrow échantillonnage des noeuds
$\mathcal{A} = \cup_{p=1..P} A_p$	Partition de \mathcal{V} appelée Parcellisation	P : nombre de parcelles dans la partition $(surf_A)_{A \in \mathcal{A}}$
$G_{\mathcal{A},bin}^0$	Graphe de connexion binaire entre les $\{A_p\}_{p=1..P}$	Distribution des connexions : <u>random regular</u> (aléatoire de degré fixe) contraint spatialement
$C_{\mathcal{A},bin}^0$	Matrice de connectivité binaire (matrice d'adjacence de $G_{\mathcal{A},bin}^0$)	idem $G_{\mathcal{A},bin}^0$
$C_{\mathcal{A},density}^0$	Matrice des densités de connexion $C_{\mathcal{A},density}^0(p, q) =$ nombre de tracts par unité de surface de la connexion entre A_p et A_q	Distribution des densités de connexion : <u>uniforme</u> exponentielle décroissante
$C_{\mathcal{A}}^0$	Matrice de connectivité $C_{\mathcal{A}}^0(p, q) =$ nombre de tracts de la connexion entre A_p et A_q	influencé par l'aire de chaque parcelle
$\{D_{A_p}\}_{p=1..P}$	Distribution des tracts parmi les points de chaque parcelle	- uniforme - aléatoire - <u>mélange</u> : uniforme (poids γ sans unité) et gaussienne centrée au centroïde (σ , étendue en mm)
G_V^0	Graphe de connexion entre les points de V	$\{D_{A_p}\}_{p=1..P}$

FIG. 4.2 – Principaux objets simulés et paramètres associés. Les paramètres choisis sont soulignés.

4.2.1 Construction du sujet *modèle*

La simulation des données pour un sujet s'effectue en 6 étapes, résumées dans la figure 4.3 qui sont les suivantes :

1. Définition et parcellisation de la surface corticale,
2. Construction de la matrice de connectivité binaire entre parcelles : représentant l'existence ou non d'une connexion entre les parcelles,
3. Construction de la matrice des densités de connexion entre les parcelles, dans laquelle des forces de connexion sont attribuées,
4. Construction de la matrice de connectivité entre parcelles représentant non pas des densités de connexions mais un nombre de tracts,
5. Modélisation des distributions locales des connexions de chaque parcelle, qui fixe, pour chaque point d'une parcelle sa probabilité à appartenir à une connexion ;
6. Construction de la matrice de connectivité des points du cortex à partir des données précédentes.

Nous allons détailler chacune de ces étapes dans la suite.

4.2.1.1 Les parcelles et leur support

Construction de la surface corticale La surface du cortex sera simulée par une sphère S , maillage triangulaire uniforme de V points v_i , d'une surface quelconque (300 cm² par exemple).

On notera plus précisément $S = (\mathcal{V}, \mathcal{E}_{spatial})$ avec :

$$\begin{cases} \mathcal{V} = (v_i)_{i=1..V} : \text{l'ensemble des noeuds ou points de } S, \\ \mathcal{E}_{spatial} : \text{les relations spatiales entre les noeuds (polygones).} \end{cases} \quad (4.1)$$

Lors de cette étape, un premier maillage triangulaire uniforme a été créé à partir de Blender ¹. Ensuite, une conversion dans le format du logiciel Brainvisa² est effectuée. Ce genre de maillage peut maintenant être construit entièrement sous Brainvisa.

Etape 1. Parcellisation de la surface corticale Ensuite on parcellise cette surface en $P = 30$ régions uniformément réparties d'aire environ 10 cm² chacune (à peu près égale à l'aire des parcelles qu'on attendrait dans les données réelles). Un algorithme de Voronoï centroïdal permet de générer ce partitionnement de la sphère (cf. Chapitre 3.2.1), qu'on notera $\mathcal{A}^0 = (A_p)_{p=1..P}$. Passons maintenant à la construction des connexions, au niveau des parcelles précédemment créées, puis au niveau des points du maillage.

4.2.1.2 Les connexions des parcelles : Graphe $G_{\mathcal{A},bin}^0$

Etape 2. Matrice de connectivité binaire $C_{\mathcal{A},bin}^0$ La construction du graphe de connexions binaire $G_{\mathcal{A},bin}^0$ entre les parcelles peut s'effectuer

¹<http://www.blender.org/>

²<http://brainvisa.info>

de différentes façons. Notons $C_{A,bin}^0$ la *matrice d'adjacence* du graphe (cf. B pour une définition) ou *matrice de connectivité*. On peut vouloir générer un graphe où les connexions sont complètement *aléatoires*, mais vouloir fixer le *degré* (cf. définition au B) des noeuds du graphe. On peut aussi vouloir rajouter des *contraintes spatiales*. Cela peut s'effectuer en jouant sur la probabilité d'existence d'une connexion entre deux parcelles $P(C_{A,bin}^0(A_p, A_q) = 1)$ qui peut dépendre de la distance entre les deux parcelles concernées. Ceci permet par exemple de privilégier une longueur de connexion plutôt qu'une autre.

Etape 3. Matrice des densités de connexion $C_{A,density}^0$ Une fois que l'on a ce graphe binaire indiquant la présence ou non d'une connexion entre deux parcelles, on ajoute un poids à ces connexions, appelé **densité de connexion**. Pour faire le parallèle avec le connectome cérébral humain, pour chaque couple de parcelles A et B, la présence d'une connexion équivaut à la présence d'un faisceau de tracts les reliant. On peut alors définir la **densité de connexion** par le nombre de tracts par unité de surface de cette connexion avec :

$$C_{A,density}^0(A, B) = \frac{2n_{tracts}}{surf_A + surf_B} \quad (4.2)$$

où n_{tracts} est le nombre de tracts du faisceau, $surf_A$ et $surf_B$ étant les aires des parcelles A et B respectivement. Cette densité de connexion est similaire à celle définie par Hagmann et al. (2008) et rappelée dans l'équation 1.1 mais ici il n'y a pas de régularisation par rapport à la longueur des tracts.

L'unité des poids est alors en nombre de tracts par mm^2 . Pour simplifier, on prend la même valeur comme *densité de connexion* pour toutes les connexions : distribution uniforme des densités de connexion. Mais on pourrait prendre une distribution exponentielle par exemple.

Hagmann et al. (2008) trouve dans son étude sur une parcellisation du cortex en 998 régions des densités de connexion allant de 2,8% à 3%, ce qui revient en nombre de tracts, à un maximum de 600 tracts par connexion (en tenant compte de la régularisation par rapport à la longueur des tracts, effectuée d'après la formule 1.1). Ainsi, des régions d'aire moyenne $1,5 cm^2$ ont des connexions de 600 tracts. Cela correspond à une densité de connexion de 4 tracts par mm^2 (d'après la formule 4.2).

Etape 4. Matrice de connectivité C_A^0 A partir de $C_{A,density}^0$, on calcule le nombre de tracts associé à chaque connexion à partir de l'équation 4.2. On appelle G_A^0 le graphe associé.

4.2.1.3 Les connexions entre les points du maillage

Chaque noeud du graphe G_A^0 est une parcelle de la surface du cortex simulée S , qui peut être décrite par la liste \mathcal{V} des *points* de S qu'elle contient. Ainsi, pour chaque connexion entre deux parcelles A et B de G_A^0 , on peut remplir $G_{\mathcal{V}}^0$, graphe de connexions entre les points de S , décrivant la répartition des tracts au sein de ces parcelles, pour le faisceau associé à cette connexion.

Etape 5. Distribution locale des connexions d'une parcelle D_A Pour chaque parcelle A , on cherche à fixer la probabilité $P(v \in \text{Connexions}_A) = D_A(v)$ qu'un de ces points v fasse partie d'une connexion de cette parcelle. On a choisi ici un mélange de deux distributions : une gaussienne fonction de la distance au centroïde c_A (centre de gravité "géodésique") de chaque parcelle A , et une distribution uniforme (sur A) avec comme poids γ .

$\forall A \in \mathcal{A}, \forall v \in \mathcal{V}$, on a :

$$D_A(v) = (1 - \gamma) \text{Gauss}_A(v) + \gamma \text{Unif}_A(v)$$

$$\text{avec} \begin{cases} \text{Gauss}_A^{\text{init}}(v) = \frac{\text{surf}_v}{\text{surf}_A} \exp\left(\frac{-\text{dist}(v, c_A)^2}{2\sigma^2}\right) \\ Z_A = \sum_{v \in \mathcal{A}} \text{gauss}_A^{\text{init}}(v) \\ \text{Gauss}_A(v) = \frac{1}{Z_A} \text{gauss}_A^{\text{init}}(v) \\ \text{Unif}_A(v) = \frac{1}{|A|} \end{cases} \quad (4.3)$$

où surf_v est l'aire des polygones entourant le point v , surf_A l'aire de A , c_A le centroïde de A , et $\text{dist}(v, c_A)$ la distance géodésique entre v et c_A .

On peut ensuite jouer sur le mélange avec γ pour favoriser une distribution plutôt qu'une autre, ainsi que sur σ pour contrôler l'étendue spatiale de la gaussienne. La figure 4.3 montre des exemples de telles distributions pour trois parcelles, avec $\sigma = 10 \text{ mm}$ et $\gamma = 0.5$.

Etape 6. Matrice de connectivité des points C_V^0 Dans le graphe de connectivité des parcelles G_A^0 , chaque connexion entre deux noeuds ou parcelles A et B correspond à $C_A^0(A, B)$ tracts. Pour chacun de ces tracts, on peut alors décider quelles extrémités (points de S) il aura dans A et B . Pour cela, pour chaque connexion, on tire n_{tracts} points parmi chaque distribution discrète D_A et D_B associée à A et à B et on met à jour le graphe G_V^0 en conséquence. Voilà la procédure plus détaillée :

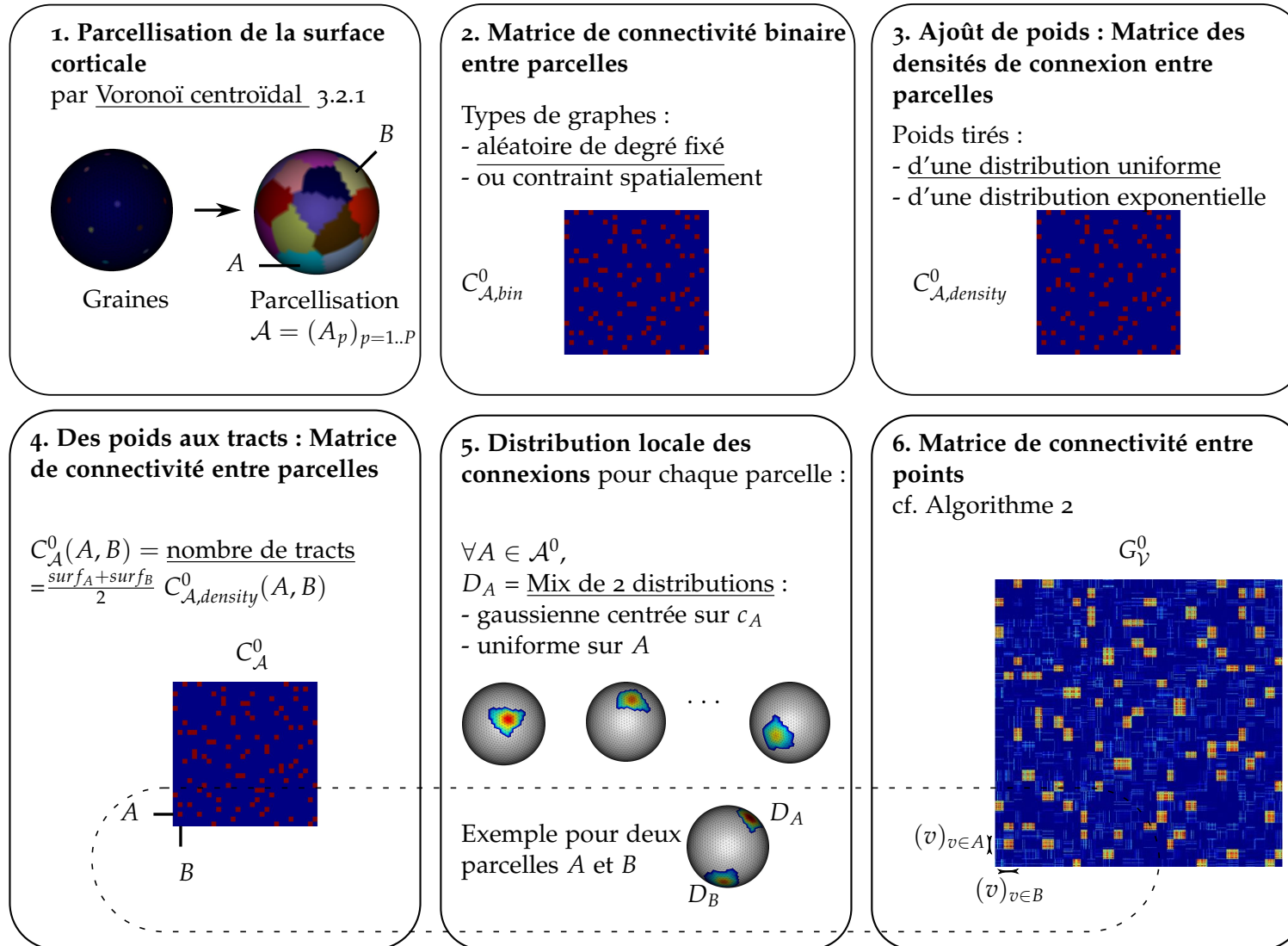
Algorithme 2: Calcul de la matrice de connectivité des points C_V^0

```

Construire  $C_V^0$ , matrice nulle de taille  $(V, V)$ 
for all  $(A, B) \in \mathcal{A}^2$  tels que  $C_A^0(A, B) \neq 0$  do
   $n_{\text{tracts}} = C_A^0(A, B)$ 
   $\text{extrémités}_A =$  vecteur de  $n_{\text{tracts}}$  points de  $A$  tiré parmi la distribution
   $D_A$ 
   $\text{extrémités}_B =$  vecteur de  $n_{\text{tracts}}$  points de  $B$  tiré parmi la distribution
   $D_B$ 
  for  $f = 1$  to  $n_{\text{tracts}}$  do
     $v_A = \text{extrémités}_A(f)$ 
     $v_B = \text{extrémités}_B(f)$ 
     $C_V^0(v_A, v_B) + = 1$ 
     $C_V^0(v_B, v_A) + = 1$ 
  end for
end for

```

FIG. 4.3 – Les différentes étapes de la simulation d'un sujet.



La figure 4.3 montre l'exemple d'une telle matrice. On peut aussi lisser géodésiquement le graphe ainsi construit avec un noyau gaussien.

4.2.2 Construction d'une base de sujets

On va maintenant expliquer la construction d'une base de N sujets à partir du sujet *modèle* (noté *sujet 0*). La création de chaque sujet s'effectue suivant les mêmes étapes que celle de la création du sujet *modèle*. Simplement, certaines données du sujet *modèle* vont être recopiées à l'identique ou légèrement modifiées, pour garantir une correspondance entre les sujets. La manière de modifier les données du *modèle* influencera la variabilité inter-sujet obtenue.

Ci-dessous, voici la liste des différentes étapes pour la création d'un sujet s , ($s \in \{1, \dots, N\}$). Les étapes où l'on modifie les données sont soulignées. Un code couleur pour les **données spécifiques au sujet** et les **données partagées entre les sujets** est ajouté à cela.

1. Parcellisation de la surface corticale $S^s = S^0 = (\mathcal{V}, \mathcal{E}_{spatial})$, \mathcal{A}^s
2. Création de la matrice binaire des connexions entre parcelles, $C_{\mathcal{A},bin}^s = C_{\mathcal{A},bin}^0$
3. Ajouter des forces de connexions : Matrice des densités de connexion entre parcelles, $C_{\mathcal{A},density}^s = C_{\mathcal{A},density}^0$
4. Conversion des poids en tracts : Matrice de connectivité des parcelles, $\bar{C}_{\mathcal{A}}^s$
5. Création des distributions locales des connexions des parcelles, $(D_{\mathcal{A}})_{\mathcal{A} \in \mathcal{A}}$
6. Création de la matrice de connectivité des points, $C_{\mathcal{V}}^s$

Etape 1. Parcellisation du cortex : réplication imparfaite de \mathcal{A}^0 A partir du partitionnement du sujet *modèle* $\mathcal{A}^0 = \cup_{p=1..P} A_p$ (avec $P = 30$ parcelles ici), on crée un nouveau partitionnement \mathcal{A}^s pour chaque sujet, de la façon suivante :

- Pour chaque parcelle A_p de \mathcal{A}^0 et le centroïde c_{A_p} associé :
Tirage aléatoire d'un point $c_p \in A_p$ de \mathcal{V} parmi les voisins de c_{A_p} (dans une "boule" géodésique de rayon d_{max} en mm)
- Nouveau Voronoï à partir des points $(c_p)_{p=1..P}$
- On a ainsi un nouveau partitionnement $\mathcal{A}^s = \cup_{p=1..P} A_p^s$

La Fig.4.4 montre dix parcellisations correspondant à dix sujets ainsi créés. Le paramètre sur lequel on pourra jouer est d_{max} , pour régler la distance du déplacement du centroïde par rapport à sa position initiale. On pourrait rajouter de nouvelles parcelles, ou en enlever. Cela n'a pas été fait pour simplifier la simulation mais ceci est envisageable.

Etapes 2-3. Matrices des connexions et des densités de connexion entre les parcelles On ne fait que recopier les matrices du sujet *modèle* pour ces deux étapes. Pour rajouter de la variabilité inter-sujet, on pourrait rajouter ou enlever des connexions entre parcelles dans la matrice des connexions $C_{\mathcal{A},bin}^s$ ou changer la densité des connexions dans $C_{\mathcal{A},density}^s$. Ceci n'est pas fait ici pour des raisons de simplicité.

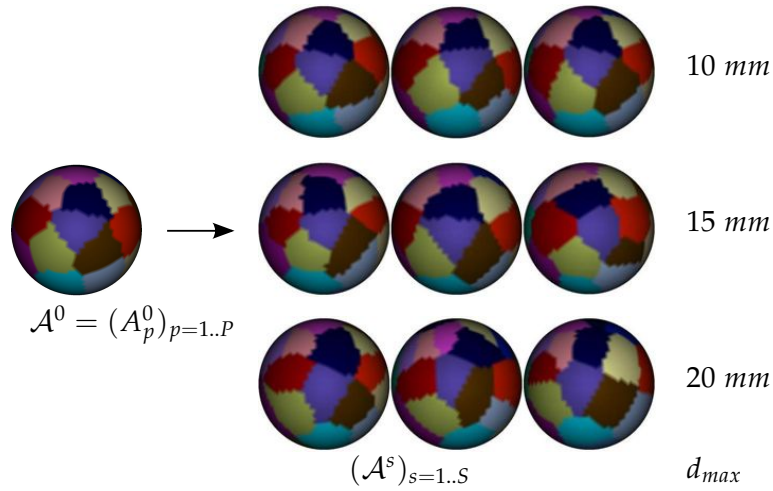


FIG. 4.4 – Données simulées. Parcellisation modèle et ensemble de sujets dérivés pour différents types de variabilité (paramètre d_{max} représentant le décalage maximum des centroïdes de chaque parcelle par rapport au sujet modèle).

Etapes 4-6. De la Matrice de connectivité des parcelles vers celle des points Pour ces étapes, on procède exactement comme pour le sujet *modèle*, à ceci près que la parcellisation a changé ($\mathcal{A}^s \neq \mathcal{A}^0$), et donc qu'on obtiendra des données différentes pour chaque sujet :

- Matrice de connectivité des parcelles $C_{\mathcal{A}}^s$: le nombre de tracts calculé pour chaque connexion dépend de l'aire des parcelles mises en jeu, qui diffère du sujet *modèle* ;
- Distribution locale des connexions des parcelles $(D_A)_{A \in \mathcal{A}}$: dépend des points associés à chaque parcelle de \mathcal{A}^s
- Matrice de connectivité des points $C_{\mathcal{V}}^s$: même remarque que pour l'étape précédente.

Résumé des hypothèses et variabilité interindividuelle résultante Considérons un ensemble de N sujets ainsi créés. Les hypothèses faites pour modéliser la variabilité inter-sujets sont les suivantes (données **invariantes** ou **non**) :

- Tous les sujets ont le **même nombre de parcelles**
- La **position** et la **taille des parcelles** varient : en pratique à partir d'un décalage des centroïdes de chaque parcelle entraînant parfois des changements de topologie dans l'agencement spatial des parcelles les unes par rapport aux autres ;
- Les **matrices de connexion et des densités de connexion des parcelles** sont les mêmes pour tous les sujets : $\forall s \in \{1, N\}$ on a : $C_{\mathcal{A}, bin}^s = C_{\mathcal{A}, bin}^0$ et $C_{\mathcal{A}, density}^s = C_{\mathcal{A}, density}^0$.
- La **matrice de connectivité des points** $C_{\mathcal{V}}^s$, est différente pour chaque sujet, reflétant le décalage spatial de chaque parcelle à travers les individus : $\forall i \in \{1, I\}$ les parcelles $(A_p^s)_{s=1..N}$ ne se superposent pas exactement mais ont un décalage reflétant le décalage de leur centroïde par rapport à la parcelle A_p^0 .

4.3 DONNÉES UTILISÉES POUR LA VALIDATION

Les données décrites ici permettront de comparer les méthodes présentées dans la suite. Ce travail sera effectué au chapitre 7. Nous allons décrire ici les caractéristiques des jeux de données simulés. Nous avons ainsi construit plusieurs bases de sujets à travers lesquelles nous pourrions évaluer le comportement des différentes méthodes en fonction de **deux critères** : la **densité des connexions** et la **variabilité interindividuelle**.

A partir d'un sujet modèle, nous avons ainsi créé des bases de sujets avec différents niveaux de variabilité inter-individuelle. Pour chaque base de sujets, les données de connectivité seront simulées à partir du sujet modèle pour différentes densités de connexion.

4.3.1 Choix des paramètres

Un résumé des paramètres choisis est présenté dans le tableau de la figure 4.5. Dans la suite nous utiliserons le code couleur suivant pour les paramètres décrits :

- Les paramètres identiques pour toutes les bases ,
- Les paramètres spécifiques à une base ,
- Les paramètres variant au sein d'une base .

4.3.1.1 Création du sujet *modèle*

Création de la surface corticale S La surface du cortex sera simulée par une sphère S , maillage triangulaire uniforme de $V = 2562$ points v_i d'une surface $surf_S = 300 \text{ cm}^2$. Cela représente une surface environ 3 fois plus petite que celle d'un hémisphère cérébral et permet ainsi de réduire la dimension des données à traiter.

Etape 1. Parcellisation de la surface corticale La surface S est subdivisée en $p = 30$ parcelles $\mathcal{A} = \cup_{p=1..p} A_p$ uniformément distribuées. Nous avons choisi ce nombre pour que les parcelles aient en moyenne une aire de 10 cm^2 .

Etape 2. Matrice de connectivité binaire entre les parcelles : $C_{\mathcal{A},bin}^0$ Cette matrice est construite de manière à former un **graphe régulier aléatoire**. Pour ce faire, j'ai utilisé des outils de génération de graphes aléatoires du package python networkx³. La fonction `random_regular_graph(d, p)` crée un graphe aléatoire de p noeuds de degré fixé d (Steger et Wormald 1999, Kim et Vu 2003). J'ai choisi ici $p = 30$ et $d = 1$ ou 3 connexions par parcelle.

³<http://networkx.lanl.gov/>

Etape 3. Matrice des densités de connexion $C_{\mathcal{A},density}^0$ J'ai choisi une distribution uniforme des densités de connexion : la même densité est attribuée à toutes les connexions. La valeur de cette densité sera nommée *densiteDeTracts*. Dans la partie validation 7 je ferai varier ce paramètre de 0,2 à 10 *tracts/mm²*.

Etape 4. Matrice de connectivité $C_{\mathcal{A}}^0$ Il n'y a pas de paramètre dans cette étape : rappelons que l'on a : $C_{\mathcal{A}}^0(A, B) = n_{tracts} = \frac{surf_A + surf_B}{2} C_{\mathcal{A},density}^0(A, B)$.

Etape 5. Distribution locale des connexions d'une parcelle D_A J'ai pris ici comme distribution la mixture de deux distributions suivant l'équation 4.3 précédente avec pour rappel :

$$D_A(v) = (1 - \gamma)Gauss_A(v) + \gamma Unif_A(v), \text{ où }$$

- $Gauss_A$ est une gaussienne fonction de la distance du point v au centroïde de A (avec comme paramètre la variance σ)
- $Unif_A$ la distribution uniforme sur A .

J'ai choisi dans un premier temps $\sigma = 5 \text{ mm}$ et $\gamma = 0.9$, privilégiant ainsi la distribution uniforme par rapport à la distribution gaussienne.

Etape 6. Matrice de connectivité des points $C_{\mathcal{V}}^0$ Il n'y a dans cette étape pas de paramètre supplémentaire. Le graphe est construit en suivant la procédure décrite précédemment, aboutissant à une matrice de taille (2562, 2562).

4.3.1.2 Création de différents jeux de données

A partir de ce sujet *modèle*, j'ai créé différentes bases de 20 nouveaux sujets en suivant la procédure décrite précédemment (cf. 4.2.2). J'ai ainsi construit au total 5 bases de 20 sujets avec pour chacune une valeur différente pour régler le décalage des centroïdes des sujets par rapport au sujet *modèle* : $d_{max} = 10, 15, 20, 30$ et 40 mm . Pour chacune de ces bases, les données de connectivité associées aux parcellisations sont générées pour 6 valeurs de densité de connexion allant de 0.2 à 1 tract par mm^2 , ceci pour chacun des sujets. Au total, l'ensemble de ces bases de sujets correspond ainsi à la création de 600 connectomes distincts.

	Etape	Paramètres
Sujet modèle	Création du cortex	$V = 2562$ points maillage triangulaire uniforme $surf_S = 300 \text{ cm}^2$
	Etape 1 : Parcellisation de la surface corticale	$p = 30$ parcelles uniformément distribuées
	Etape 2 : Matrice de connectivité binaire entre les parcelles	graphe régulier aléatoire de degré $d = 1$ ou 3
	Etape 3 : Matrice des densités de connexion	distribution uniforme des densités de connexion $densiteDeTracts$ variant de $0,2$ à 10 tracts/mm^2
	Etape 4 : Matrice de connectivité	aucun paramètre
	Etape 5 : Distribution locale des connexions d'une parcelle	$D_A(v) = (1 - \gamma)Gauss_A(v) + \gamma Unif_A(v)$ $\sigma = 5 \text{ mm}$ et $\gamma = 0.9$ ou 0.5
	Etape 6 : Matrice de connectivité des points	aucun paramètre
Base de sujets dérivés	Création du cortex	pour $N = 20$ sujets
	Etape 1 : Parcellisation de la surface corticale	$d_{max} = 10$ à 40 mm
	Etapes 2 à 6	paramètres du sujet <i>modèle</i>

FIG. 4.5 – Choix et variation des paramètres des données simulées.

4.3.2 Deux bases de données pour illustrer les méthodes

En plus des bases précédentes, deux bases de sujets ont été construites pour illustrer le comportement des différentes méthodes.

4.3.2.1 Cas optimal

Une première base de données est construite dans un cas optimal défini par :

1. une densité de tracts élevée (10 tracts/mm^2),
2. une distribution locale des connexions d'une parcelle centrée autour du centroïde avec $\sigma = 5 \text{ mm}$ et $\gamma = 0.5$,
3. un total de 3 connexions par parcelle,
4. une variabilité interindividuelle peu élevée avec $d_{max} = 10 \text{ mm}$.

Ainsi un sujet modèle est construit (avec les 3 premiers paramètres précédents). Une base de données de 10 sujets est ensuite générée à partir de ce modèle selon la variabilité interindividuelle précédente. La figure 4.6(a) présente les parcellisations de 10 sujets de base. Cette base de données sera appelée dans la suite $B_{optimale}$.

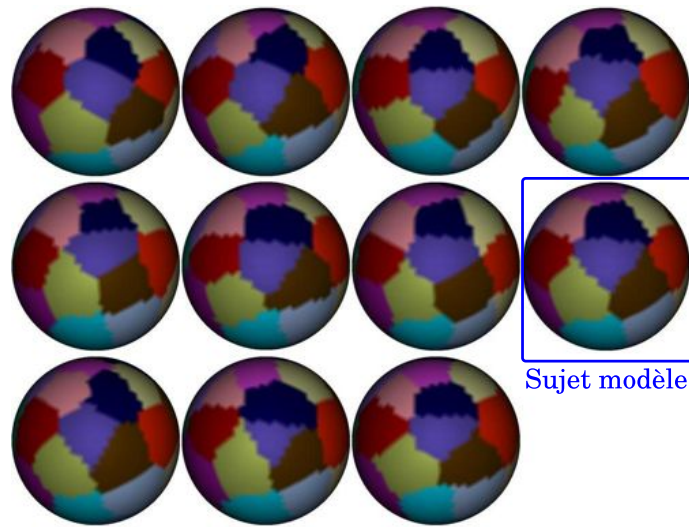
4.3.2.2 Cas extrême

La deuxième base de données représente le cas extrême défini par :

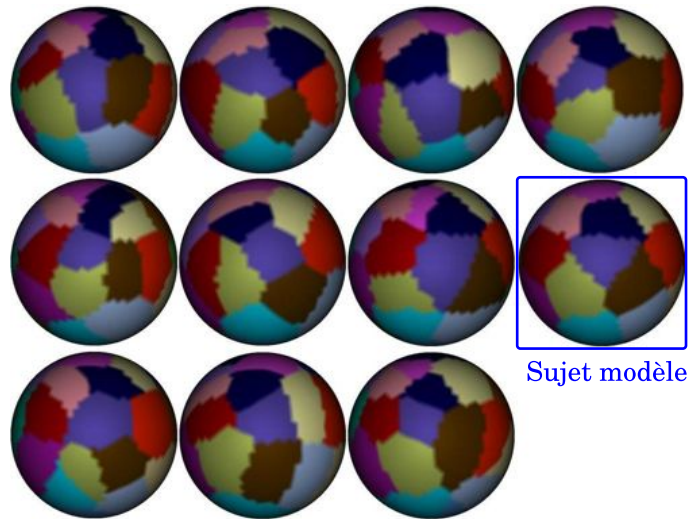
1. une densité de tracts faible (0.3 tracts/mm^2),

2. une distribution locale des connexions d'une parcelles moins gaussienne que précédemment avec $\sigma = 5 \text{ mm}$ et $\gamma = 0.9$ (contribution de la distribution uniforme).
3. une seule connexion par parcelle : la matrice de connectivité binaire entre les parcelles est un graphe aléatoire régulier de degré 1.
4. une variabilité du double de celle du cas optimal précédent ($d_{\max} = 20 \text{ mm}$).

Cette base de données sera appelée dans la suite B_{extreme} . Les parcellisations des différents sujets sont présentées dans la figure 4.6(b).



(a) Cas optimal



(b) Cas extrême

FIG. 4.6 – Les deux bases de données construites pour illustrer les méthodes. Cette figure présente les parcellisations des 10 sujets de chacune des bases, ainsi que celle du sujet modèle dont elles sont dérivées. Dans les sujets dérivés, le centroïde de chaque parcelle a été décalé par rapport au sujet modèle d'une distance maximum de $d = 10 \text{ mm}$ pour le cas optimal et de 20 mm pour le cas extrême.

4.4 CONCLUSION

Dans ce chapitre, nous avons présenté une manière de simuler les **connectomes de plusieurs bases de sujets en contrôlant la variabilité inter-individuelle au niveau des noeuds** du connectome de ces sujets **ainsi que la densité des connexions associées**. Pour construire une base de sujets donnée, un sujet modèle est tout d'abord fabriqué avec : une parcellisation du cortex et les données de connectivité au niveau des parcelles ainsi qu'au niveau des points de la surface corticale constituant chacune des parcelles.

En ce qui concerne les données de connectivité, nous avons ainsi mis en place une procédure permettant de les contrôler à deux niveaux. A l'échelle des parcelles, nous contrôlons le nombre de connexions de ces parcelles (degré des noeuds du graphe de connectivité entre les parcelles) ainsi que la densité des connexions associées. A l'échelle des points de la surface corticale, nous modélisons aussi la répartition spatiale des connexions par un mélange de deux distributions gaussienne et uniforme simulant ainsi le caractère resserré ou non des fibres vers le centre des faisceaux.

Des nouveaux sujets sont ensuite construits à partir du sujet modèle. La variabilité spatiale dans la position et la forme des parcelles constituant les éléments de base de ces connectomes est effectuée en introduisant un décalage entre les centroïdes des parcelles de chaque sujet et ceux des parcelles du sujet modèle. La nouvelle position du centroïde est tirée aléatoirement dans un disque géodésique autour du centroïde correspondant du sujet modèle. Le rayon de ce disque contrôle ainsi la variabilité maximale autorisée. Un développement futur intéressant serait de tirer cette position non pas aléatoirement mais à partir d'une distribution de probabilité plus appropriée bien sûr.

Finalement, nous avons simulé 5 bases de 20 sujets chacune associées à des variabilités inter-individuelles croissantes. Pour chacune de ces bases, les données de connectivité ont été générées pour 6 valeurs de densité de connexion. Nous avons donc généré un total de **600 connectomes distincts**.

Nous pourrions ainsi étudier par la suite le comportement des méthodes présentées en fonction de ces degrés de variabilité et de densité de connexion. Nous avons de plus généré **deux bases exemples** correspondant à deux cas opposés : un *cas optimal* où la variabilité inter-individuelle est faible et la densité des connexions élevée, et le cas inverse appelé *cas extrême*. Ces deux bases de sujets nous serviront par la suite à illustrer les méthodes présentées, notamment celles qui réduisent la dimension des données de connectivité et qui seront abordées dans le chapitre 5 suivant.

RÉDUCTION DES PROFILS DE CONNECTIVITÉ

LE profil de connectivité d'un point du cortex représente ses connexions. A chaque type de matrice de connectivité défini précédemment correspond un nouveau profil de connectivité. En effet, un profil de connectivité n'est rien d'autre qu'une ligne ou une colonne de la matrice associée. Dans le cas de la connectivité anatomique, cette matrice est symétrique car la direction des connexions n'est pas accessible.

L'enjeu est de comparer de tels profils, pour chaque individu, ou pour un groupe de sujets. L'un des premiers problèmes est la dimension des données à manipuler. En effet, si l'on considère un échantillonnage du cortex en 40 000 points, nous aurons alors 40 000 profils de taille 40 000 par sujet. L'une des principales difficultés est de trouver un espace approprié pour réduire la dimension des profils. Trouver une mesure de distance adéquate entre les profils est de plus important.

Dans ce chapitre, nous présenterons dans une première partie une solution pour résumer les profils de connectivité à l'échelle de l'individu, puis au niveau du groupe, pour ensuite présenter quelques mesures de distances en les comparant.

SOMMAIRE

6.1	DONNÉES ET HYPOTHÈSES	133
6.2	MÉTHODE	135
6.2.1	Division du domaine	135
6.2.2	Traitement de chaque cellule	137
6.2.2.1	1. Réduction de dimension des profils de connectivité	137
6.2.2.2	2. Clustering des profils réduits	137
6.2.2.3	Etapes 3. Sélection des parcelles valides et 4. Mise à jour de la parcellisation finale	138
6.2.3	Test d'arrêt	138
6.3	ESSAIS SUR LES DONNÉES SIMULÉES	138
6.3.1	Comparaison des parcellisations	138
6.3.2	Cas optimal	139
6.3.3	Cas extrême	139
6.4	ESSAIS SUR LES DONNÉES RÉELLES	141
6.4.1	Résultats	141
6.5	CONCLUSION	144

5.1 INTRODUCTION

Différents profils de connectivité

Dans la littérature, deux principaux types de profils de connectivité sont exploités (cf. chapitre 1) :

- le **profil complet** d'une région (voxel, point ou groupe de tels éléments), qui représente ses forces de connexion vers l'ensemble des éléments d'un ensemble d'intérêt (volume de la matière blanche, cerveau entier ou interface gris blanc appelée ici surface corticale),
- le **profil réduit d'une région vers des régions d'intérêt** : régions définies à partir d'a priori (contraintes sulco-gyrales, tracés manuels s'inspirant d'études chez le primate par exemple).

Dans ce chapitre, nous allons considérer les données d'un morceau connexe de la surface corticale que nous appellerons patch (cf. Notations 5.1).

Nous proposerons différents moyens de construire automatiquement des profils de connectivité intermédiaires à ces deux types précédents à partir des données de connectivité elles-mêmes :

1. Nous présenterons ainsi une façon d'identifier automatiquement des régions d'intérêt à partir des données de connectivité du patch, ceci pour chaque sujet ou au niveau d'un groupe de sujets. Cet ensemble de régions d'intérêt permettra alors de mieux visualiser la connectivité du patch en question en utilisant le mode de visualisation interactif présenté dans la partie 3.4.2.
2. Dans un deuxième temps, nous utiliserons la méthode d'Analyse en Composantes Principales pour projeter les profils de connectivité dans un espace de dimension plus faible que celui des profils complets.

Différentes distances entre ces profils

Nous présenterons dans un second temps différentes mesures de distance ou de similarité possibles entre les profils de connectivité. Dans la littérature, la principale mesure de distance utilisée est la cross-corrélation entre les profils complets (approche aveugle). Nous aborderons notamment une autre distance, la distance de Hausdorff, qui permet de prendre en compte l'organisation spatiales des connexions présentes dans chaque profil de connectivité lors du calcul de la distance et qui peut être une mesure intéressante et plus appropriée dans certains cas. Nous illustrerons ensuite à l'aide des données simulées de parcellisation et de connectivité (cf. Chapitre 4) comment la réduction des profils de connectivité joue sur la similarité entre les profils d'une parcellisation.

NOTATIONS

En reprennant les notation précédentes (cf. Annexe Notations B), on représentera la surface corticale par $S = (\mathcal{V}, \mathcal{E}_{spatial})$ avec :

$$\begin{cases} \mathcal{V} = (v_i)_{i=1..V} : \text{l'ensemble des noeuds ou points de } S, \\ \mathcal{E}_{spatial} : \text{les relations spatiales entre les noeuds (polygones).} \end{cases}$$

Considérons que l'on veuille parcelliser une portion ou **patch** de la surface corticale (cf Fig. 5.1 pour des exemples) à partir du profil de connectivité de chacun de ces points. L'ensemble des points du cortex constituant ce **patch** est noté \mathcal{V}_{patch} , de cardinal P . Nous considérons que ce sont les P premiers points $\{v_1, ..v_P\}$ de \mathcal{V} pour simplifier les notations. Pour chaque point $v_i \in \mathcal{V}_{patch}$ du patch, on peut définir un **profil de connectivité** p_i vers l'ensemble des points de la surface corticale. On obtient ainsi une matrice de connectivité $\mathbf{X} \in M_{(P,V)}(\mathbb{R}^+)$:

$$\mathbf{X} = \begin{pmatrix} v_1 & \dots & v_V \\ \dots & p_1 & \dots \\ & \vdots & \\ \dots & p_P & \dots \end{pmatrix} \begin{matrix} v_1 \\ \\ v_P \end{matrix}$$

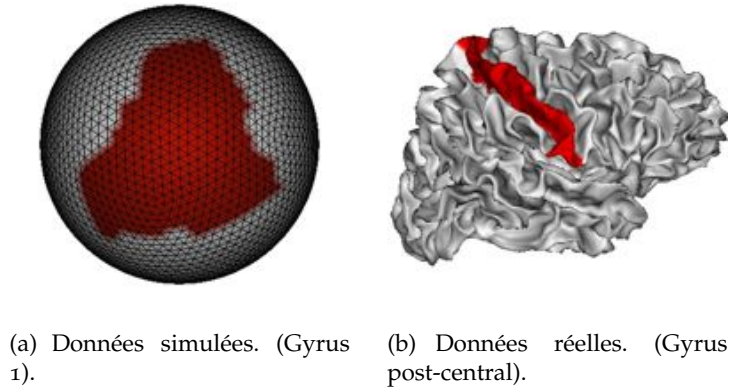


FIG. 5.1 – Exemple de définition de patches sur la surface corticale.

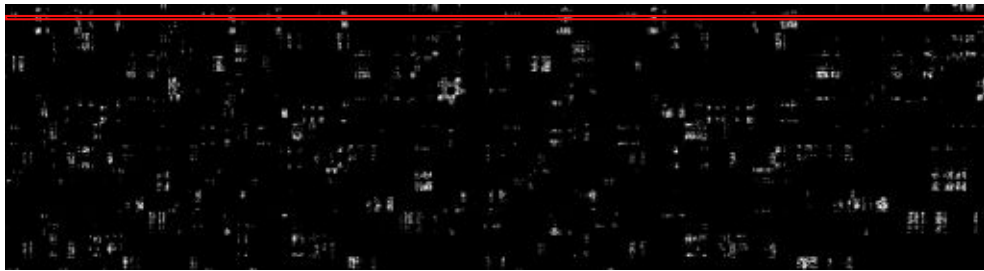


FIG. 5.2 – Exemple de matrice de connectivité d'un patch de la surface corticale. Le profil de connectivité d'un point de ce patch est entouré en rouge.

5.2 IDENTIFICATION AUTOMATIQUE DE RÉGIONS D'INTÉRÊT

5.2.1 Méthode Intra-sujet

Nous proposons ici une approche intermédiaire aux deux types de profils de connectivité précédents, qui identifie un ensemble de régions d'intérêts (ROIs) à partir des données de connectivité elles-mêmes. L'idée est simple. Un organigramme est présenté sur la Figure 5.3.

On fait ici l'hypothèse que ce n'est pas utile de différencier des points du cortex (dans le profil de connectivité) qui appartiennent au même faisceau. On veut regrouper le profil de connectivité vers un ensemble de ROIs associées aux régions connectées au *patch*.

On réalise pour ce faire une réduction de dimension des données de connectivité qui se découpe en trois étapes :

1. Calcul du profil de connectivité (total) du patch,
2. Identification de régions d'intérêt (ROIs) fortement connectées au patch : ligne de partage des eaux à partir du profil précédent,
3. Calcul de la matrice de connectivité réduite du patch en fonction de ces ROIs.

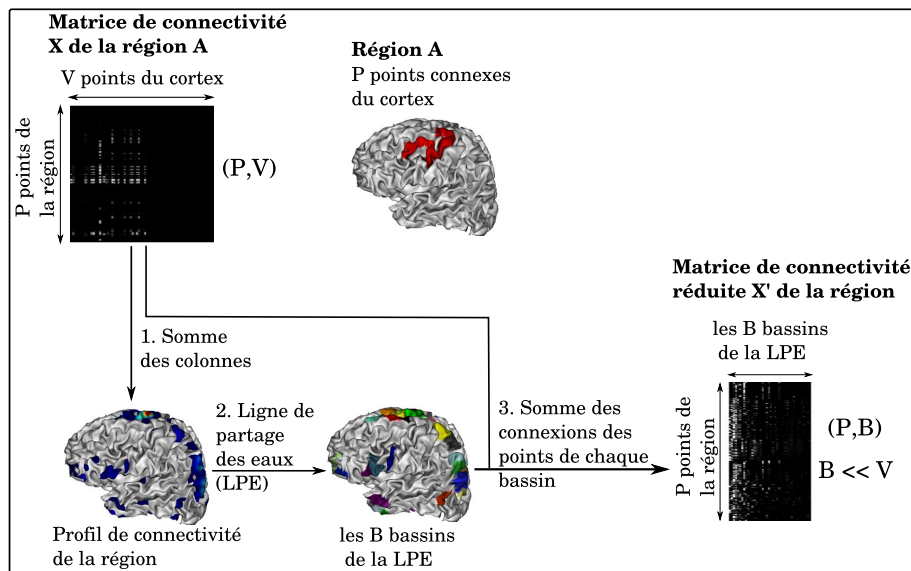
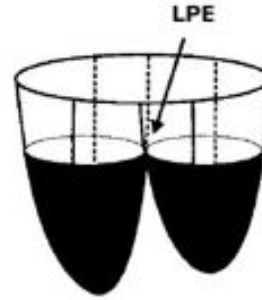


FIG. 5.3 – Les différentes étapes de la réduction de dimension intra-sujet.

Étape 1 : Calcul du profil de connectivité du patch On peut calculer la *connectivité totale* ou profil de connectivité du *patch*, qui n'est rien d'autre que la somme des profils de connectivité de chacun de ses points (points appartenant au patch). Ce profil de connectivité se représente sous la forme d'une texture sur la surface corticale (cf. Fig. 5.3.1), présentant des maximas correspondant à l'arrivée des faisceaux au niveau du cortex. Les endroits du cortex avec de fortes connexions anatomiques avec le patch apparaissent en rouge sur cette figure. En pratique, ce profil de connectivité est lissé géodésiquement à la surface pour faciliter le déroulement de l'étape suivante ($\sigma = 5 \text{ mm}$).

FIG. 5.4 – Principe de la ligne de partage des eaux. Une image en niveau de gris peut être considérée comme un relief topographique (où l'altitude est définie à partir des niveaux de gris). Ce relief est inondé à partir de ses minima locaux. Lorsque deux bassins se rencontrent, une ligne de partage des eaux (LPE) est créée.



Etape 2 : Identification des ROIs fortement connectées au patch Pour extraire ces régions connectées au patch, on effectue ensuite une **ligne de partage des eaux** sur cette texture de connectivité totale (cf. Fig. 5.4 pour le principe de l'algorithme de ligne de partage des eaux). Ceci est fait de manière géodésique à la surface corticale (cf. Fig. 5.3.2), et aboutit au découpage du cortex en B bassins, chaque bassin correspondant à un maximum local de la connectivité. Les bassins obtenus par cet algorithme correspondent aux régions d'intérêt ou *ROIs* connectées au patch. La figure 5.5 montre les ROIs obtenues et les faisceaux de tracts associés sur un exemple. En pratique un processus de regroupement des bassins trouvés par la ligne de partage des eaux a été mis en place : les bassins de taille trop petite (contenant 10 points ici) ou de profondeur trop faible sont fusionnés au bassin le plus proche. De plus le profil de connectivité du patch est seuillé avant la ligne de partage des eaux à partir d'un seuil que l'on appellera $seuil_{LPE}$. Ce paramètre est à régler en fonction du patch considéré. Dans le cas du gyrus post-central par exemple, le seuil choisi correspond à la médiane des valeurs d'intensité de ce profil. En effet, dans ce cas, la moitié des connexions non nulles restent très proches de zéro.



FIG. 5.5 – Bassins LPE et tracts associés obtenus à partir de la région A définie dans la figure 5.3.

Etape 3 : Calcul de la matrice de connectivité réduite du patch Les profils de connectivité des points du *patch* sont alors réduits par rapport à ces ROIs. Plus précisément, on obtient le profil de connectivité réduit d'un point à partir de son profil de connectivité complet (vers tous les points du cortex) et de la segmentation du cortex en ROIs précédente. Pour chacune des B ROIs, les connexions des points correspondants sont

sommées pour aboutir à une matrice de connectivité réduite \mathbf{X}' de taille (P, B) :

$$\mathbf{X}' = \begin{pmatrix} R_1 & \dots & R_B \\ \dots & p'_1 & \dots \\ & \vdots & \\ \dots & p'_P & \dots \end{pmatrix} \begin{matrix} v_1 \\ \\ v_P \end{matrix} \quad (5.1)$$

avec : $\begin{cases} \forall (i, j) \in \llbracket 1, P \rrbracket \times \llbracket 1, B \rrbracket : \mathbf{X}'(i, j) = \sum_{k, v_k \in R_j} \mathbf{X}(i, k) \\ \text{La ligne } \mathbf{X}'(i) \text{ est le profil réduit du point } v_i, \text{ noté } p'_i \\ \forall j \in \llbracket 1, B \rrbracket, R_j \text{ est la } j^{\text{ème}} \text{ ROI issue de la ligne de partage des eaux.} \end{cases}$

5.2.2 Extension de la méthode à un groupe de sujets par moyennage

Cette réduction des profils prend toute son importance quand on passe à l'analyse d'un groupe de N sujets. Si l'on trouve de telles ROIs à l'échelle du groupe, on aura alors une base de décomposition commune des profils de connectivité.

Le plus facile est de faire le même traitement que précédemment, mais en remplaçant la connectivité totale du patch par sa connectivité moyennée à travers le groupe de sujets. On peut ensuite effectuer une ligne de partage des eaux sur cette nouvelle carte surfacique, et identifier des ROIs. La Figure 5.6 présente la méthode. On utilise pour cela l'appariement surfacique donné par freeSurfer, décrit dans la partie 2.3.2.2.

Plus précisément, la méthode se déroule en quatre étapes :

1. Sujet par sujet : Calcul du profil de connectivité du patch,
2. A travers les sujets : Moyenne des profils de connectivité précédents,
3. Sur le profil moyenné : Identification des ROIs fortement connectées au patch,
4. Sujet par sujet : Calcul de la matrice de connectivité réduite du patch en fonction de ces ROIs.

Etape 1 : Sujet par sujet : Calcul du profil de connectivité du patch

Cette étape est la même que pour la méthode intra-sujet précédente, à ceci près que les profils de connectivité calculés au niveau des points du patch doivent être calculés à partir de la surface corticale extraite par freeSurfer qui présente une correspondance point à point entre les points à travers les sujets. Ainsi, chaque sujet a P profils de connectivité (un pour chaque point du patch, le patch étant le "même" pour tous les sujets) qui représente la force de connexion des points du patch vers les V points de la surface corticale, P et V ne variant pas à travers les sujets. Le profil du patch est alors un vecteur de dimension V . On obtient ainsi N profils correspondant aux différents sujets.

Etape 2 : A travers les sujets : Moyenne des profils précédents Un moyennage est ensuite effectué entre ces N profils à travers les N sujets.

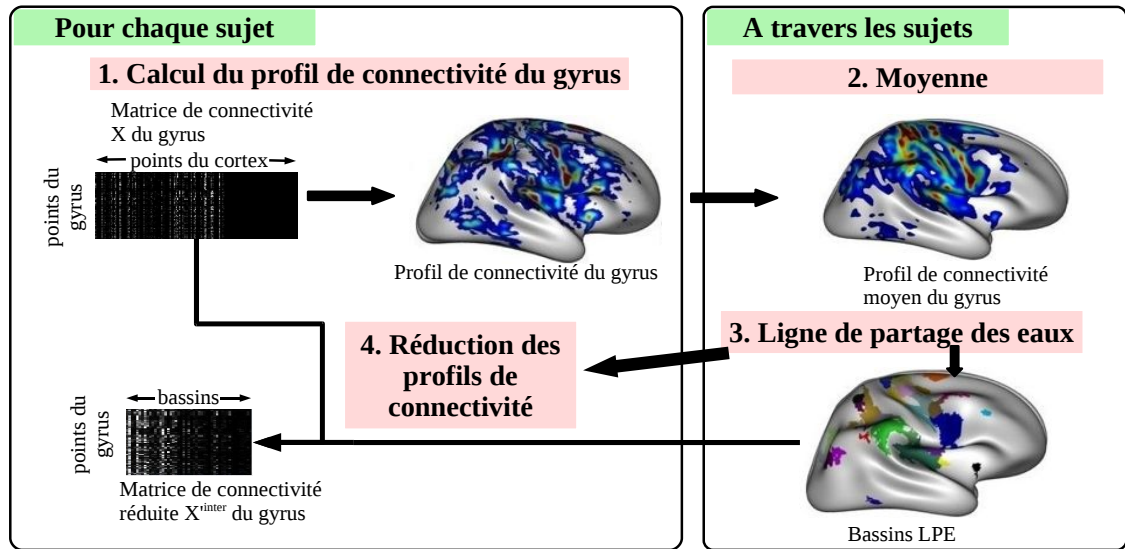


FIG. 5.6 – Les différentes étapes de la réduction de dimension inter-sujets.

On obtient ainsi un profil de connectivité moyen du patch, vecteur de dimension V . Ce profil est lissé pour faciliter l'étape suivante, ceci étant fait géodésiquement (avec $\sigma = 5 \text{ mm}$).

Etape 3 : Sur le profil moyenné : Identification des ROIs connectées au patch Comme dans la méthode intra-sujet précédente, on effectue une ligne de partage des eaux sur le profil du patch (mais moyenné cette fois-ci), aboutissant ainsi à une segmentation de la surface corticale en B bassins. Préalablement, le profil de connectivité moyen du patch est seuillé.

Etape 4 : Sujet par sujet : Calcul de la matrice de connectivité réduite du patch en fonction de ces ROIs La même procédure qu'à l'étape 3. de la méthode intra-sujet est appliquée (cf.5.2.1).

5.2.3 Données Simulées : cas optimal




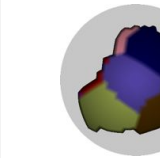
Nous nous plaçons dans un premier temps dans le cas optimal (cf. partie Simulation 4.3.2.1 et Fig. 5.7), avec :

- une densité de tracts élevée (10 tracts/mm^2),
- la distribution locale des connexions d'une parcelle centrée autour du centroïde avec $\sigma = 5 \text{ mm}$ et $\gamma = 0.5$,
- Chaque parcelle a **trois connexions** vers trois parcelles de la surface corticale.

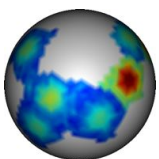
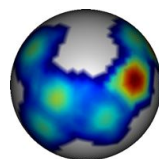


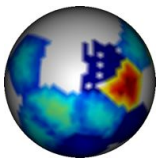
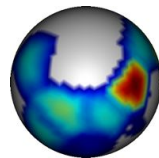


Intra-sujet Nous considérons une parcellisation en gyri du sujet *modèle* faite de telle sorte que les frontières des "gyri" et des parcelles coïncident, chaque gyrus couvrant en moyenne 5 parcelles. Nous présentons ici l'analyse effectuée sur un de ces gyri, que l'on nommera gyrus 1. Une autre configuration est aussi envisagée, c'est celle où le gyrus considéré est légèrement décalé par rapport aux parcelles du sujet. Ce cas est constitué

en prenant un sujet de la base simulée $B_{optimale}$ (cf. 4.3) où les centroïdes des parcelles du sujet sont décalées par rapport à ceux du sujet *modèle* de 10 mm au maximum. La figure 5.7(a) présente ces deux configurations.

1. Déroulement de la méthode de réduction de dimension des profils de connectivité par identification de ROIs (cf. 5.2) : La figure 5.7(b) montre les résultats obtenus pour les différentes étapes : le profil de connectivité du gyrus calculé (avant et après lissage géodésique gaussien) ainsi que les bassins de la ligne de partage des eaux obtenus. Dans la première configuration, tout se déroule pour le mieux. Les bassins correspondent en effet aux parcelles initiales. Le deuxième cas se comporte de façon similaire. Le seuillage du profil moyen élimine les connexions des parcelles peu représentées dans le gyrus 1 (parcelles rose et rouge).

Modèle T_1		Sujet 1 $\in B_{optimale}$	
Parcellisation	Gyri	Parcellisation	Parcelles du gyrus
			

(a) Données simulées. Le Gyrus 1 se trouve au centre. Sur le sujet modèle les frontières des parcelles bleu foncé, beige et marron coïncident avec celles du gyrus. Sur le sujet 1, elles ne coïncident pas et d'autres parcelles (rose et rouge) sont présentes dans le gyrus.

	Profils de connectivité moyens	Après lissage et seuillage	Bassins LPE	Parcellisation
Modèle				
Sujet 1				

(b) Différentes étapes pour le gyrus 1. Vue centrée autour de la région présentant de fortes connexions avec le Gyrus 1. Une rotation est effectuée par rapport à la vue 5.7(a) ci-dessus.

FIG. 5.7 – Analyse d'un gyrus de la base simulée dans deux cas de figures : - Cas du sujet modèle (les frontières du gyrus et celles des parcelles coïncident), - Sujet avec une parcellisation décalée par rapport au gyrus. 5.7(a) Données simulées dans les deux cas : avec la parcellisation en 30 parcelles et la parcellisation en gyri ; 5.7(b) Les premières étapes de la méthode de réduction de dimension dans les deux cas de figure.

2. Comparaison avant/après réduction des profils : On constate qu'avec les profils réduits, il y a une meilleure séparation des parcelles.

Le calcul de la largeur de silhouette de chaque point du gyrus permet de le remarquer (ce critère est appelé “silhouette avg width” en anglais). En effet, la largeur de silhouette moyenne est un nombre réel entre 0 et 1, d’autant plus grand que le point est proche (au sens de la distance entre les profils de connectivité) des points du cluster auquel il appartient et éloigné des autres clusters (cf. partie 6.2.2.2 pour une définition plus complète). En moyennant sur chaque parcelle cette largeur de silhouette (obtenant ainsi la “cluster average silhouette width” en anglais), on obtient un critère indiquant à quel point la parcelle est bien séparée des autres parcelles, ainsi que traduisant l’homogénéité des profils au sein de cette parcelle. La moyenne des “silhouette width” sur l’ensemble des points du gyrus est nommée “l’average silhouette width” et indique si globalement les parcelles sont bien séparées et homogènes. La figure 5.9 montre ainsi ces largeurs de silhouette pour deux cas : le cas avant réduction des profils (Fig. 5.9(a)) et le cas après réduction (Fig. 5.9(b)). Ainsi, avant réduction l’average silhouette width est de 0.33 contre 0.91 après réduction. Les silhouette width individuelles (pour chaque parcelle, et pour chaque point) sont beaucoup plus grandes pour les profils de connectivité réduits, indiquant une meilleure séparation des parcelles et une meilleure homogénéité après réduction des profils (moins de variance intra-parcelle).

Une analyse de scaling multidimensionnel métrique classique (Gower 1966) permet de représenter les points dans un espace bidimensionnel où les distances sont relativement conservées (distance entre termes de profil de connectivité). Cette représentation reflète bien la meilleure séparation des parcelles ainsi que leur homogénéité dans le cas des profils réduits (cf. Fig. 5.8).

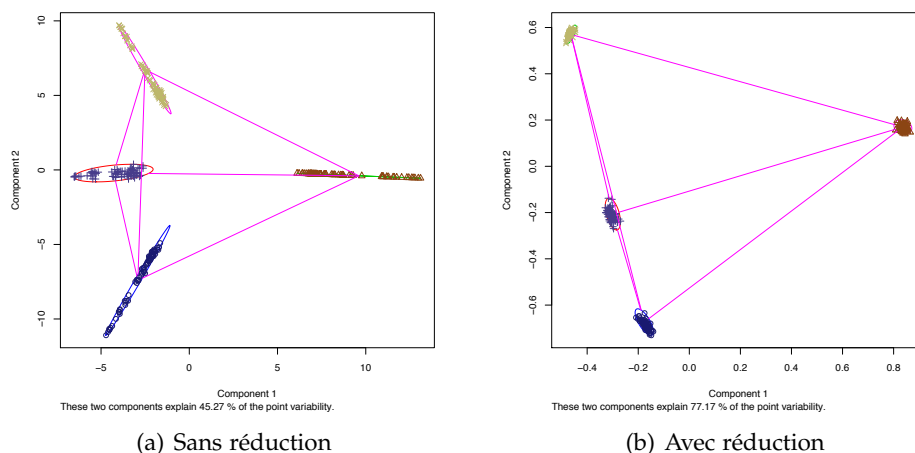
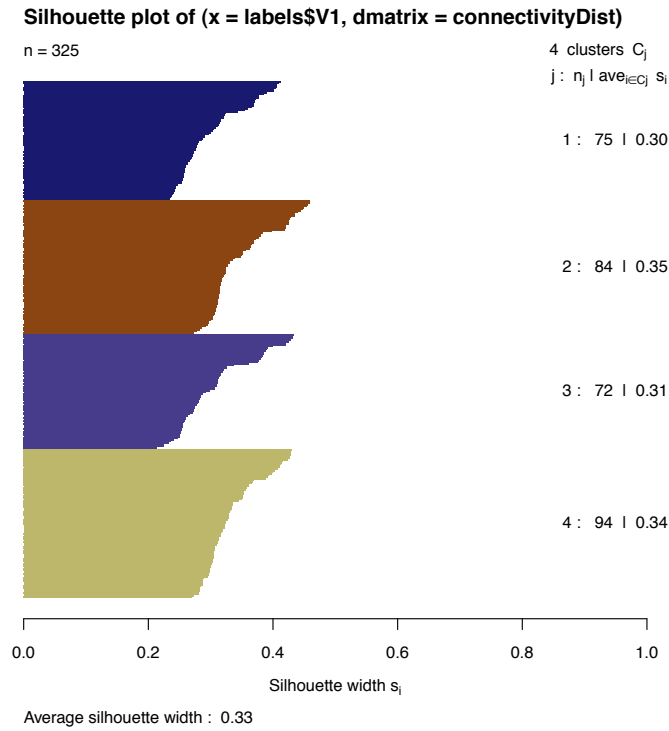
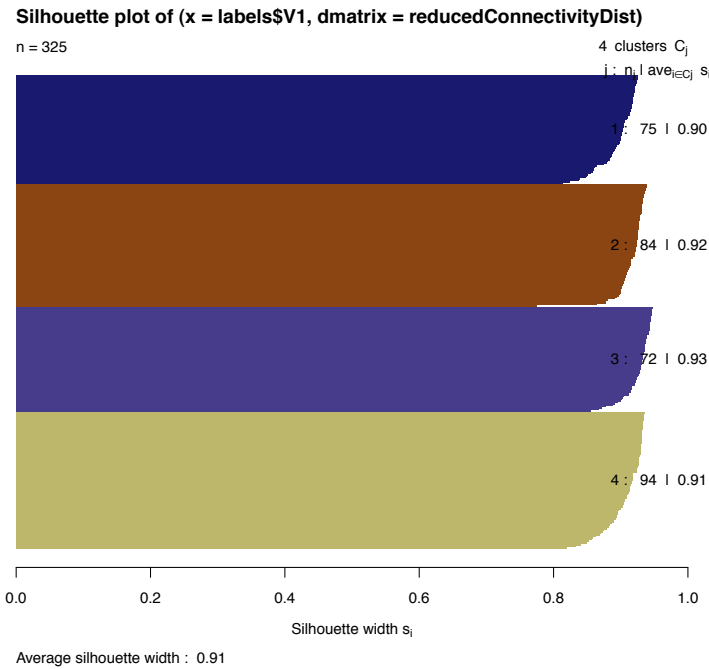


FIG. 5.8 – Cas optimal : Représentation bidimensionnelle des parcelles avant et après réduction des profils de connectivité du sujet modèle (faite par scaling multidimensionnel, utilisation de la fonction *clusplot* de R). Les parcelles ont l’air mieux séparées avec la matrice de distance euclidienne sur les profils réduits qu’avec celle construite avec les profils complets. La variance intra-parcelle est également plus faible.



(a) Sans réduction : la largeur de silhouette de chaque parcelle est inférieure à 0.34.



(b) Avec réduction : la largeur de silhouette de chaque parcelle est supérieure à 0.90.

FIG. 5.9 – Cas optimal : Silhouette avant et après réduction des profils de connectivité du sujet modèle, pour chaque point du gyrus 1 (points en ordonnée, largeur de silhouette en abscisse). Les parcelles sont mieux séparées avec la matrice de distance euclidienne sur les profils réduits qu'avec celle construite avec les profils complets.

Inter-sujets : Sujets dérivés de T_1 peu décalés : base $B_{optimale}$ Rappelons que les données simulées des 10 sujets de la base $B_{optimale}$ ont les mêmes caractéristiques que le sujet *modèle* T_1 (densité élevée de tracts, distribution locale des connexions d'une parcelle centrée autour du centroïde), avec une variabilité au niveau des parcelles de chaque sujet (les centroïdes de chaque parcelle étant décalés de 10 mm maximum par rapport au sujet *modèle*). Pour plus de détails, veuillez vous reporter à la description des données simulées 4.3. L'analyse est focalisée sur le gyrus 1 de la parcellisation en gyri du sujet *modèle* T_1 comme précédemment (cf. Fig. 5.7(a)).

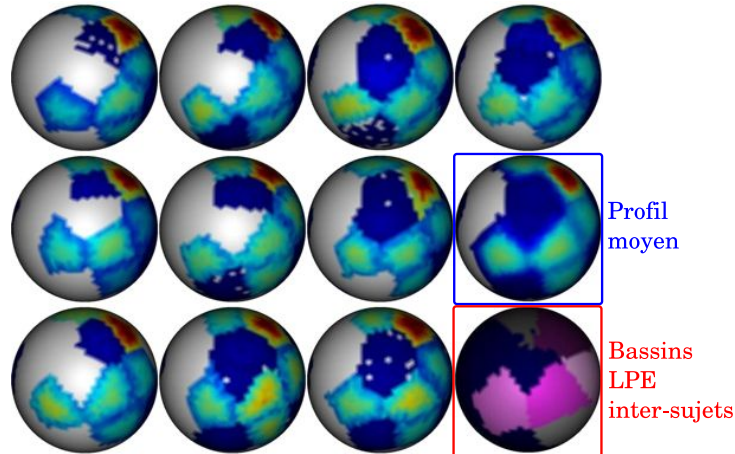
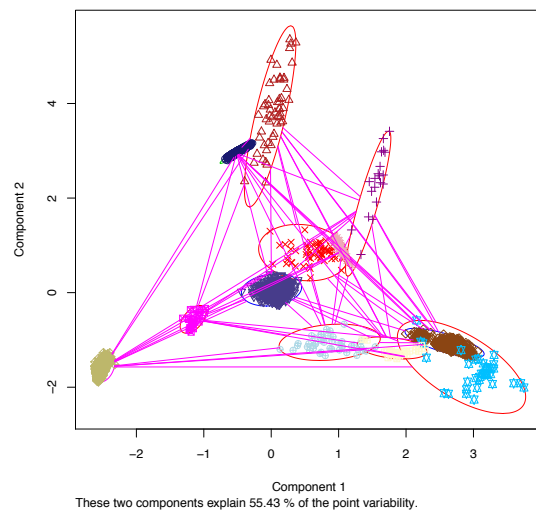


FIG. 5.10 – Base de sujets simulés $B_{optimale}$. Résultats de la réduction de dimension inter-sujets : Profils de connectivité du gyrus 1 de chacun des 10 sujets (non entourés) ainsi que le profil moyenné à travers les sujets (entouré en bleu), et les bassins issus de la ligne de partage des eaux effectuée sur ce profil (entouré en rouge).

Les figures 5.10 et 5.12 présentent les résultats obtenus : les profils de connectivité du gyrus pour chaque sujet, ainsi que les bassins obtenus par ligne de partage des eaux dans les cas intra- et inter-sujets.

FIG. 5.11 – Cas inter-sujets. Base de sujets optimale $B_{optimale}$. Projection des profils de connectivité réduits suivant les 2 premières composantes d'une ACP. Les couleurs correspondent aux parcelles du gyrus 1 (beige, violet, bleu foncé et marron pour les plus grosses parcelles).

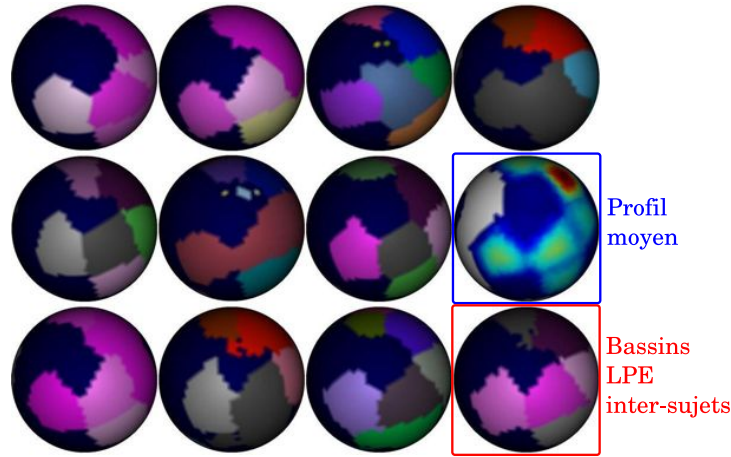


1. Cas intra-sujet : Les bassins obtenus sont plutôt cohérents, mais certains bassins regroupent des parcelles (cf. Fig. 5.12). Ce sont souvent

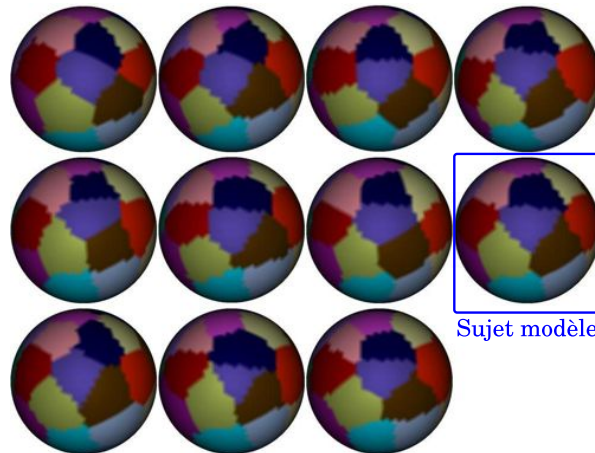
des parcelles proches, et avec des connexions faibles au gyrus qui sont regroupées.

2. Cas inter-sujets : Les bassins obtenus correspondent bien chacun à une parcelle : il n'y a pas de regroupement comme dans le cas précédent (cf. Fig. 5.12). Cela peut s'expliquer par le moyennage, qui renforce les maxima de connexions de chaque parcelle (étant donné le décalage léger des centroïdes entre les sujets, ainsi que la répartition gaussienne des connexions autour des centroïdes des parcelles).

Pour se faire une idée du pouvoir de séparation qu'il y a entre les différentes parcelles décrites par leur profil de connectivité réduit, une analyse en composantes principales (ou ACP) peut être effectuée sur ces profils. La figure 5.11 montre ainsi la projection des profils réduits suivant les deux premières composantes d'une telle analyse. Les principales parcelles du gyrus sont bien séparées (parcelles beige, marron, bleu foncé et violette).



(a) Bassins obtenus par LPE intra-sujet (non entourés) ou inter-sujets (entouré en rouge) et profil moyen du patch (entouré en bleu).



(b) Parcellisations des sujets et du sujet modèle (entouré en bleu).

FIG. 5.12 – Base de sujets simulée B_{optimale} . Réduction intra-sujet et inter-sujets. 5.12(a) Bassins obtenus par LPE intra-sujet pour chacun des sujets, profil de connectivité moyenné et bassins de la LPE inter-sujets ; 5.12(b) Parcellisations simulées des sujets ainsi que celle du sujet modèle .

5.2.4 Données Simulées : Base extrême

Nous nous plaçons maintenant dans le cas extrême conformément à la description faite dans la partie précédente (cf. partie 4.3.2.2) avec un sujet modèle ayant :

- une connexion seulement par parcelle : la matrice de connectivité binaire entre les parcelles est un graphe aléatoire régulier de degré 1,
- une densité de tracts faible ($0.3 \text{ tracts}/\text{mm}^2$),
- une distribution locale des connexions d'une parcelles est moins gaussienne (autour du centroïde) avec $\sigma = 5 \text{ mm}$ et $\gamma = 0.9$.

A partir de ce *modèle*, une base de sujets avec une variabilité assez importante ($d_{\max} = 20 \text{ mm}$) est construite : B_{extreme} . La figure 5.13 montre les parcellisations obtenues pour les différents sujets.

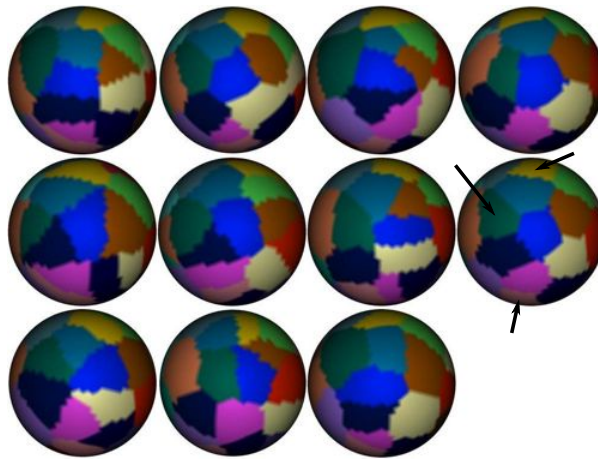


FIG. 5.13 – Parcellisations simulées de chacun des sujets de la base ainsi que celle du sujet modèle (avec les flèches qui indiquent les principales parcelles connectées au gyrus 1, connexions visibles sur les profils de connectivité de la figure 5.14).

La figure 5.14 montre les profils de connectivité de chaque sujet. Le gyrus a trois connexions principales visibles sur cette vue correspondant aux parcelles vert foncé, jaune et rose clair (indiquées par des flèches sur la parcellisation du sujet *modèle* sur la figure 5.13).

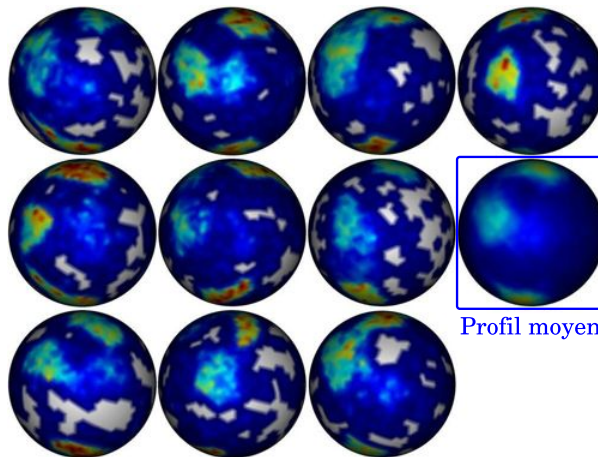


FIG. 5.14 – Cas extrême B_{extreme} . Profils de connectivité du gyrus 1 de chacun des 10 sujets (non entourés) ainsi que le profil moyenné à travers les sujets (entouré en bleu).

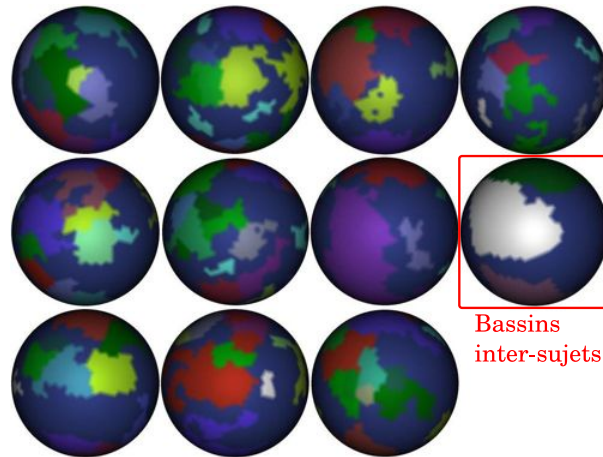


FIG. 5.15 – Cas extrême $B_{extreme}$. Bassins résultant de la ligne de partage des eaux en intra-sujet pour chacun des sujets (non entourés) ou en inter-sujets (entouré en rouge).

La figure 5.15 montre les bassins obtenus à partir de la ligne de partage des eaux intra-sujet et inter-sujet :

1. Cas intra-sujet : La plupart du temps, les bassins obtenus ne correspondent pas à des parcelles initiales, ceci s'expliquant par le peu de connexions entre les parcelles. Plusieurs maxima locaux apparaissent au sein d'une parcelle connectée et la procédure de regroupement des bassins ne les regroupe pas.
2. Cas inter-sujets : Les bassins obtenus sont plus cohérents : le moyennage à travers les sujets remplit son rôle. Chaque bassin de la ligne de partage des eaux englobe l'une des trois parcelles connectées au gyrus.

La figure 5.16 montre la projection des profils de connectivité réduits suivant les deux premières composantes principales d'une ACP (de la même façon que la figure 5.11 pour le cas optimal). La séparation entre les parcelles est moins évidente ici.

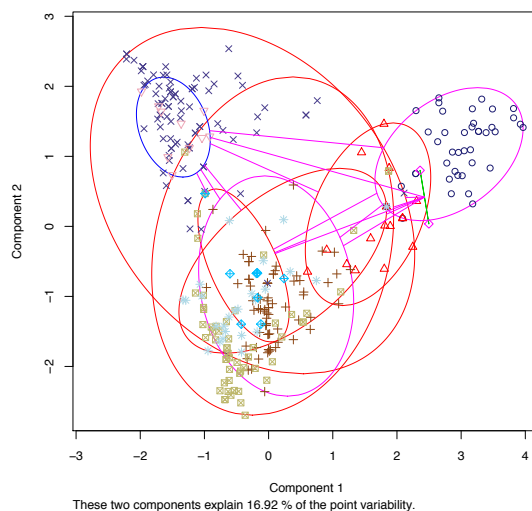


FIG. 5.16 – Cas intra-sujet, Base de sujets simulée extrême $B_{extreme}$. Projection des profils de connectivité réduits suivant les deux premières composantes d'une analyse en composantes principales. Les couleurs correspondent aux parcelles.

5.2.5 Essais sur les données réelles

Description des données Les données de connectivité anatomique utilisées ici sont celles de la deuxième base de données (cf. partie 2.4.2). En résumé, un tracking probabiliste est effectué sur les données HARDI de la base NMR. L'analyse FreeSurfer nous fournit des maillages du cortex avec une correspondance point à point à travers les sujets. Pour chaque sujet, la matrice de connectivité des points de la région d'intérêt est calculée suivant la méthode décrite à la partie 3.3. La matrice C_s d'un sujet $s \in M_{(P,V)}(\mathbb{N})$ où P et V sont respectivement le nombre de points de la région d'intérêt (patch) ou de la surface totale du cortex. Il y a alors une double correspondance entre les matrices des différents sujets. Chaque ligne correspond au profil de connectivité d'un point de la région d'intérêt, point qui se retrouve chez tous les sujets. Chaque colonne de la matrice correspond à un point de la surface corticale, qui se retrouve aussi à travers les sujets.

Gyrus Post-central

Résultats La figure 5.17 montre les profils de connectivité intra-sujet, le profil moyen du gyrus post-central à travers les sujets ainsi que les régions obtenues par ligne de partage des eaux sur le profil moyen.

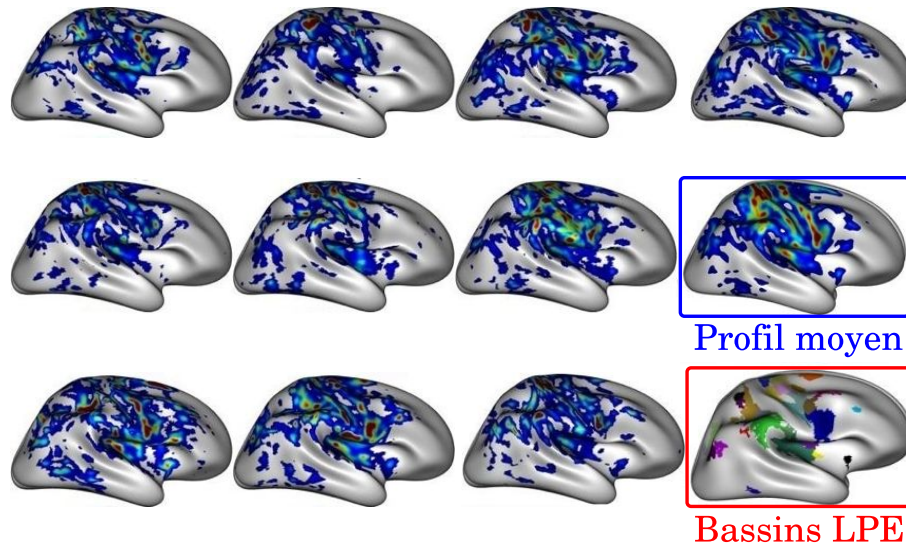


FIG. 5.17 – Profils de connectivité du gyrus post-central pour les 10 sujets (non entourés), profil moyenné à travers les individus (entouré en bleu), et bassins obtenus par la ligne de partage des eaux sur ce profil (entourés en rouge).

Remarques En regardant de plus près la *connectivité totale* individuelle d'un *patch* (profil de connectivité du patch) à travers les sujets, on constate que les ROIs extraites individuellement ne se correspondent pas parfaitement quand elles sont projetées sur la surface corticale moyenne. Il peut y avoir des décalages spatiaux non négligeables, dus par exemple aux erreurs de recalage de la méthode surfacique d'appariement utilisée, aux erreurs de tracking, ou à une variabilité inter-individuelle des faisceaux.

Le moyennage du profil à travers les sujets peut apporter une réponse dans le cas où la répartition du centre des connexions (position géodésique sur la surface) est gaussienne.

La Figure 5.18 par exemple, présente un zoom sur une partie du profil de connectivité du gyrus-post-central, pour trois sujets (5.18.A. s_1, s_2, s_3). On distingue une région de forte connexion à des positions différentes pour chaque sujet. Sur le profil moyen associé (5.18.avg), les régions précédentes se regroupent pour former une seule région. Cette région est ensuite extraite par la ligne de partage des eaux 5.18.wat. La partie 5.18.B représente deux types de profils de connectivité individuel d'un point du cortex pour les trois sujets (simplifié) :

- soit le profil vers le cortex entier p (appelé aussi profil complet),
- soit le profil p' réduit vers les ROIs déterminées précédemment.

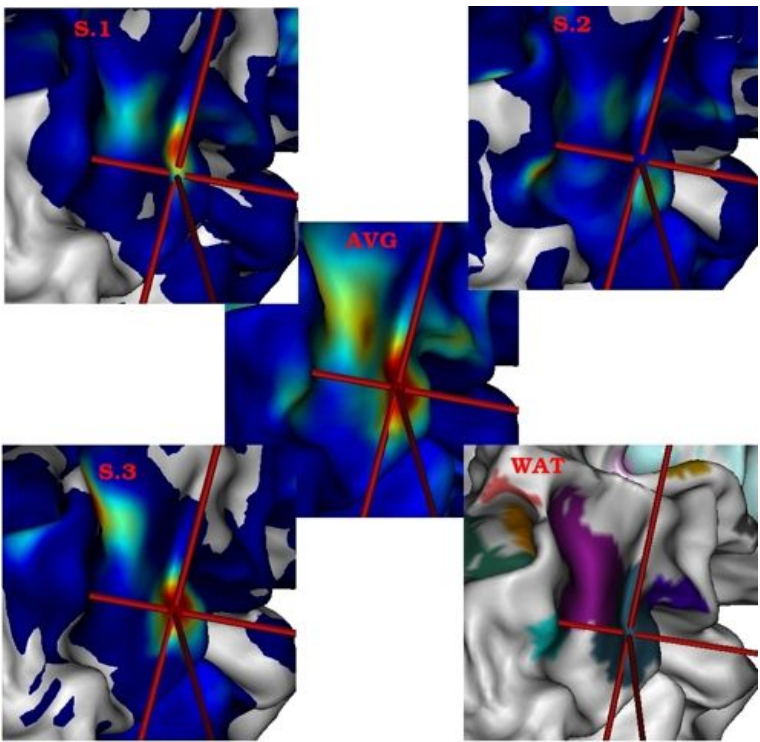
Dans l'exemple, les profils réduits $(p^s)_{s=1..3}$ des différents sujets sont identiques, alors que les profils complets $(p'^s)_{s=1..3}$ sont différents. Si la distance euclidienne est utilisée pour comparer ces profils, lors d'un clustering ultérieur, cela sera plus facile de regrouper les points des 3 sujets dans un même cluster à partir des profils réduits $(p^s)_{s=1..3}$, où la similarité entre les profils des sujets est maximale.

Région englobant le SMA et le pré-SMA Nous avons défini une région d'intérêt initiale englobant l'aire motrice supplémentaire (SMA) et l'aire prémotrice associée (pré-SMA) (cf. partie 2.3.2.3).

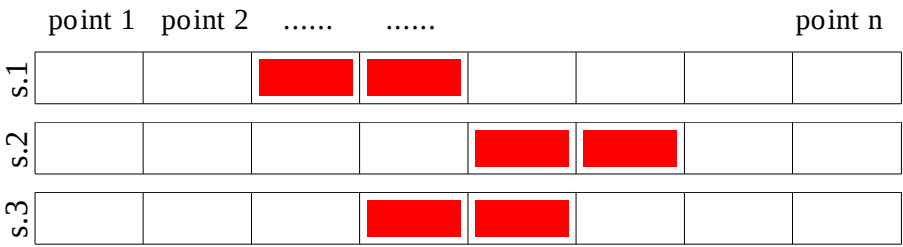
La figure 5.19 montre les profils de connectivité de cette région pour chaque sujet, ainsi que leur moyenne à travers les sujets. Les connexions retrouvées sont cohérentes avec la littérature (cf. partie 1.2.3.2). Les zones fortement connectées à la région étant le cortex moteur, le cortex frontal, le cortex pariétal médian, ainsi que la partie de "l'interface du cortex" correspondant au passage du tract cortico-spinal.

Les figures 5.19(b) et 5.19(d) montrent les résultats de la ligne de partage des eaux qui identifie un ensemble de régions d'intérêt, au niveau intra et inter-sujets. En comparaison du cas intra-sujet, les bassins trouvés en inter-sujets sont moins nombreux et plus lisses.

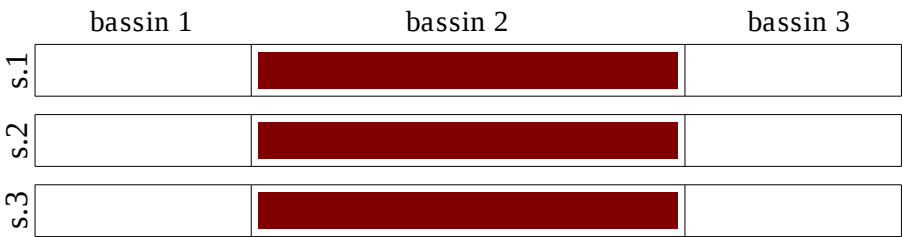
La figure 5.20 montre les profils de connectivité de chaque bassin trouvé en inter-sujets, pour deux sujets : un maillage texturé représentant la région englobant le SMA et le pré-SMA est positionné au dessus de chaque bassin. La texture représente le nombre de connexions total du bassin vers chacun des points de la région. En visualisant deux sujets, la variabilité apparaît déjà importante, même si des similarités apparaissent aussi, les connexions les moins variables étant les plus courtes (connexions vers le cortex moteur par exemple).



(a) Visualisation sur la surface corticale moyenne. s1, s2, s3 : Profils de connectivité des 3 sujets, AVG. Profil de connectivité moyenné, WAT. Les bassins associés.



(b) Exemple 2D des profils de connectivité complets des 3 sujets.



(c) Exemple 2D des profils réduits des 3 sujets selon les 3 bassins LPE.

FIG. 5.18 – Intérêt de la LPE inter-sujets. Illustration sur un zoom autour d’une partie du profil de connectivité du gyrus post-central pour 3 sujets. Les connexions entre les sujets présentent un léger décalage spatial (Fig. 5.18(a).s1-3). Les sujets ont donc des profils de connectivité complets différents (Fig. 5.18(b)). Dans le cas optimal, la LPE effectuée sur le profil moyen (Fig. 5.18(a).AVG) regroupe les connexions des différents sujets dans un même bassin (Fig. 5.18(a).WAT). Les profils réduits des 3 sujets vers ces bassins sont alors identiques dans l’exemple ci-dessus (Fig. 5.18(c)).

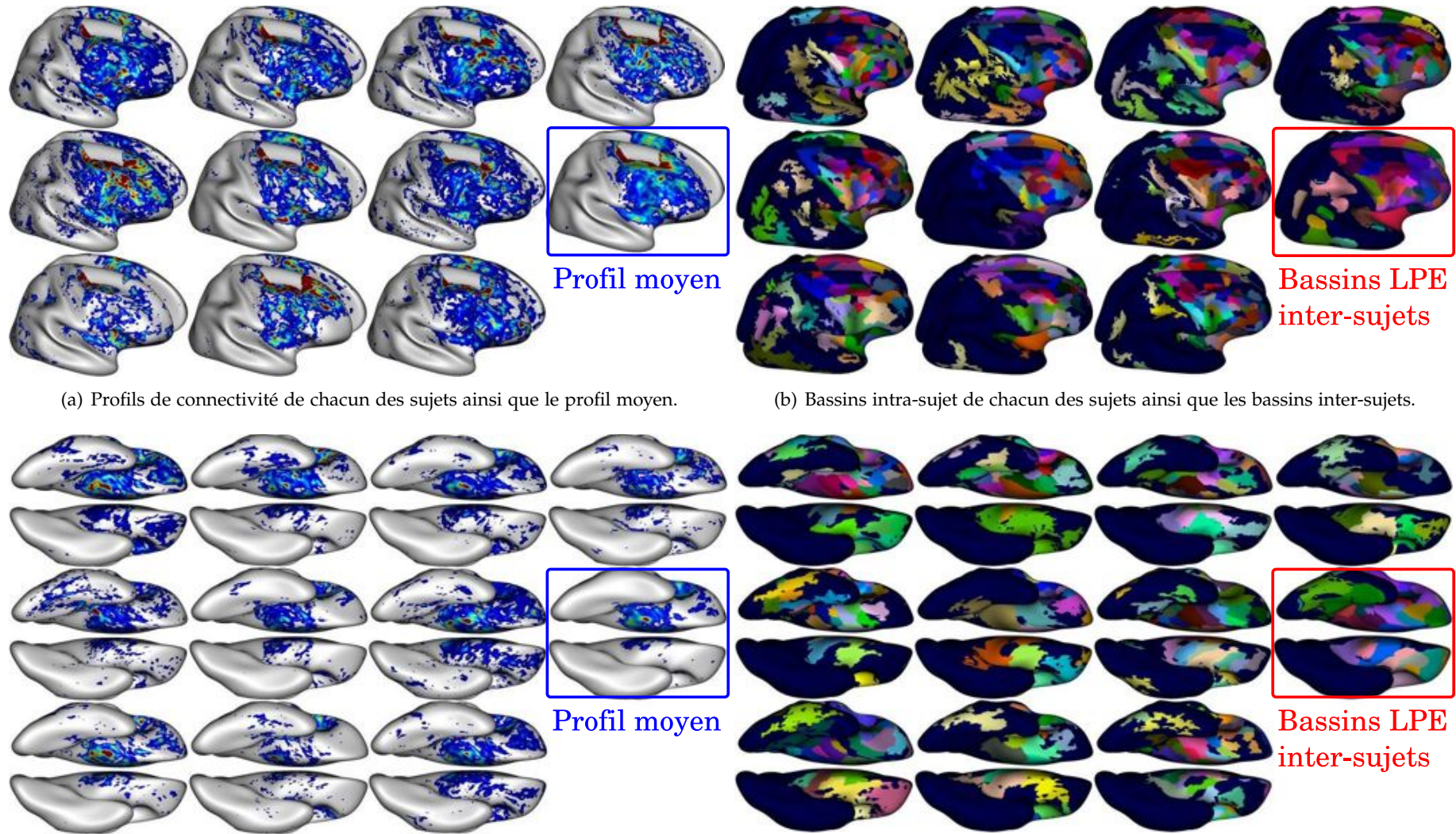
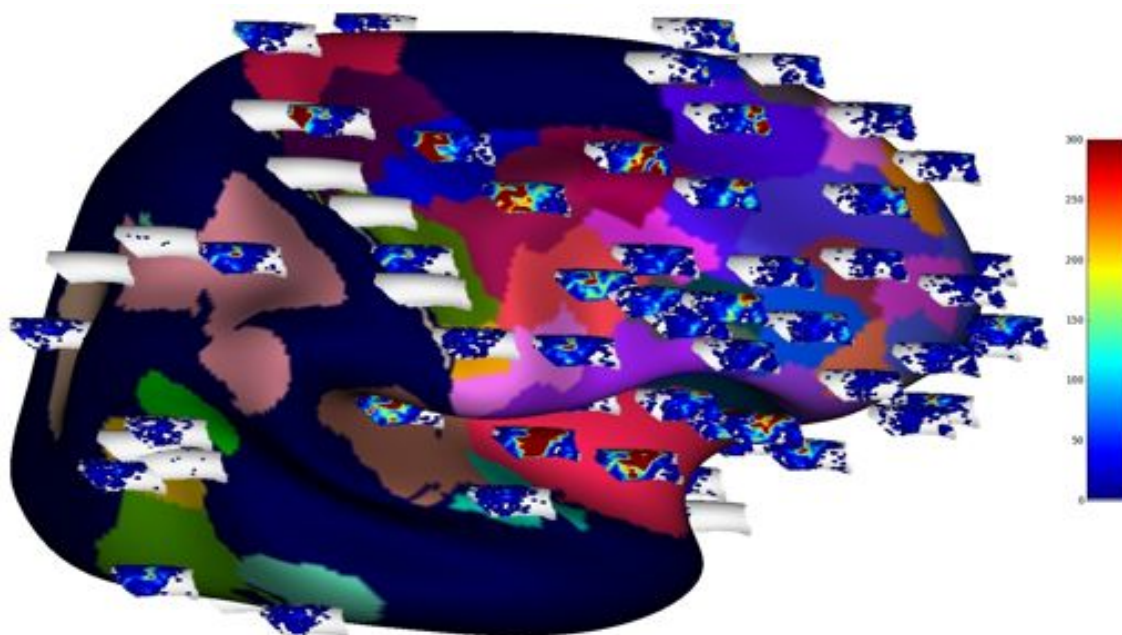
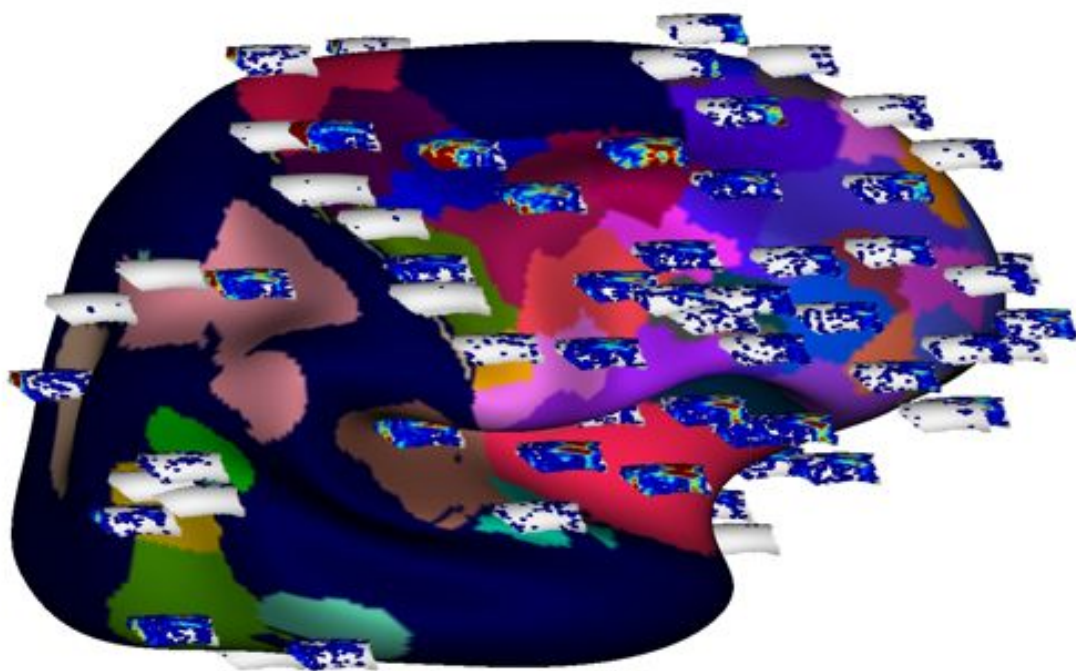


FIG. 5.19 – Réduction de dimension par identification des ensembles de régions d'intérêt en intra et inter-sujets. Exemple de la région $R_{SMA \cup \text{pré-SMA}}$ englobant le SMA et le pré-SMA. 5.19(a) Profils de connectivité. 5.19(b) Régions obtenues en intra et inter-sujets. Il y a des connexions avec le cortex moteur, le cortex frontal, ainsi que le cortex pariétal médian.



(a) Sujet01



(b) Sujet03

FIG. 5.20 – Connectivité de la région englobant le SMA et le pré-SMA pour deux sujets. Pour chaque bassin identifié au niveau inter-sujets, un modèle réduit de la région englobant le SMA et le pré-SMA est ajouté avec une texture représentant les connexions du bassin vers chaque point de la région (échelle allant du bleu au rouge foncé pour toutes les connexions de plus de 300 tracts).

5.3 ANALYSE EN COMPOSANTES PRINCIPALES

Une autre approche pour réduire les profils de connectivité sans a priori est d'utiliser une méthode de réduction de dimension classique, l'analyse en composantes principales par exemple. Dans ce cas, l'information de voisinage et l'organisation des points sur la surface corticale sous-jacente ne seront pas utilisées contrairement à la méthode précédente.

5.3.1 Principe

Nous nous plaçons exactement dans le même cas que celui défini précédemment (cf. 5.1. Notations). Nous considérons un patch de la surface corticale constitué de P points auxquels sont associés P profils de connectivité corticaux i.e vers l'ensemble du cortex (constitué de V points). Rappelons que ces profils de dimension V représentent les lignes d'une matrice de connectivité $\mathbf{X} \in M_{(P,V)}(\mathbb{R}^+)$.

L'ensemble des profils (appelés encore **observations**) sont des points de l'espace $(\mathbb{R}^+)^V$, qui forment ainsi un nuage de points dans cet espace appelé espace des **attributs**.

L'**analyse en composantes principales** consiste à trouver un sous-espace \mathcal{H} de $(\mathbb{R}^+)^V$ de dimension k dans lequel on aura la meilleure représentation du nuage de points constitué par les profils, c'est-à-dire le sous-espace sur lequel la projection du nuage de points aura la plus grande variance (Hotelling 1933).

Soit \mathbf{X} une matrice de données centrées, l'analyse en composantes principales cherche une projection linéaire \mathbf{M} qui maximise la fonction coût : $\text{trace}(\mathbf{M} \text{cov}(\mathbf{X}) \mathbf{M})$ où $\text{cov}(\mathbf{X})$ est la matrice de covariance de \mathbf{X} .

On peut montrer que cela est équivalent à trouver les k **valeurs propres de la matrice de covariance de \mathbf{X}** correspondant au k plus grandes valeurs propres.

Dans cette thèse, l'analyse en composantes principales est effectuée avec le langage R (Venables et Ripley 2002).

5.3.2 Méthode

La démarche adoptée se déroule en deux étapes :

1. Calcul des profils de connectivité vers une segmentation uniforme de la surface corticale (avec contraintes ou non),
2. Analyse en composantes principales de la matrice constituée par ces profils réduits, soit au niveau intra-sujet, soit en inter-sujets.

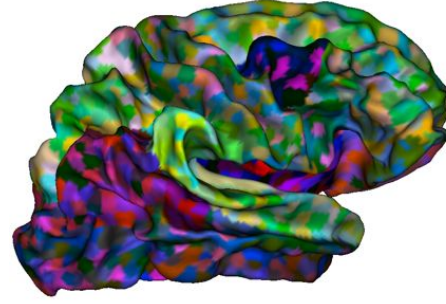
1. Première réduction de dimension à partir d'une segmentation a priori

En pratique, nous avons un problème de dimension car la matrice $\mathbf{X} \in M_{(P,V)}(\mathbb{R}^+)$ contient plus d'attributs que d'observations : P le nombre de points du patch considéré est obligatoirement inférieur à V , le nombre de points de la surface corticale complète. Pour résoudre ce problème, une première réduction des profils est effectuée.

Segmentation utilisée Dans le cas d’une approche volumique, certaines études travaillant sur les tractogrammes¹ changent de résolution et calculent les tractogrammes sur des voxels plus gros (5 mm isotropique par exemple dans l’article de Johansen-Berg et al. (2004)). Dans cette thèse, la définition d’une segmentation de Voronoï centroïdale (Du et al. 1999) de la surface du cortex remplit le même rôle.

Cette segmentation fondée sur des diagrammes de Voronoï itératifs aboutit à une parcellisation en cellules de Voronoï de tailles similaires réparties uniformément à la surface du cortex. Une description de l’algorithme et de son implémentation surfacique est détaillée dans la partie spécifique aux prétraitements surfaciques 3.2.1. Nous noterons $\{\mathcal{V}_k\}_{k=1}^K$ la partition de la surface corticale \mathcal{V} en résultant.

FIG. 5.21 – Diagramme de Voronoï centroïdal. Exemple de parcellisation en 2400 régions de la surface corticale de l’hémisphère droit. Le maillage de cette surface utilisé ici est le maillage moyen de la base de sujets de cette thèse fourni par Freesurfer (cf. 2.3.2.2).



Réduction des profils de connectivité La réduction des profils de connectivité constituant la matrice de connectivité \mathbf{X} du patch s’effectue de la même manière qu’à l’étape 3 (cf. Equation 5.1) en remplaçant les régions $\{\mathcal{R}_b\}_{b=1}^B$ par les $\{\mathcal{V}_k\}_{k=1}^K$. Cette étape aboutit donc à une matrice de connectivité réduite \mathbf{X}' du patch du type :

$$\mathbf{X}' = \begin{pmatrix} V_1 & \dots & V_K \\ \dots & p'_1 & \dots \\ & \vdots & \\ \dots & p'_p & \dots \end{pmatrix} \begin{matrix} v_1 \\ \\ v_p \end{matrix}$$

$$\text{avec : } \begin{cases} \forall (i, j) \in \llbracket 1, P \rrbracket \times \llbracket 1, K \rrbracket : \mathbf{X}'(i, j) = \sum_{k, v_k \in R_j} \mathbf{X}(i, k) \\ \text{La ligne } \mathbf{X}'(i) \text{ est le profil réduit du point } v_i, \text{ noté } p'_i \\ \forall j \in \llbracket 1, K \rrbracket, V_j \text{ est la } j^{\text{ème}} \text{ région de la segmentation utilisée.} \end{cases}$$

Normalisation des profils de connectivité Lorsque le but ultime est de parcelliser le patch en régions homogènes en terme de profils de connectivité, une étape de normalisation peut être effectuée. En effet, les profils réduits précédent représentent pour chaque point son **nombre de tracts** vers chacune des régions $\{\mathcal{V}_k\}_{k=1}^K$ précédentes. On peut vouloir clusteriser les points du patch non pas en fonction du nombre de tracts mais en fonction de la **proportion de chaque connexion** (associée à une région

¹ profils de connectivité vers l’ensemble des voxels de la matière blanche ou du cerveau

\mathcal{V}_k dans les connexions de chaque point). L'étape de normalisation a pour but de transformer le nombre de connexions en proportions.

En pratique, la **matrice X'** est d'abord **normalisée** suivant la norme $\| \cdot \|_1$. Chaque profil réduit p'_i est divisé par la somme des forces de connexion le constituant :

$$\forall i \in \llbracket 1, P \rrbracket : Z_i = \sum_{k=1}^K p'_i(k)$$

$$\forall j \in \llbracket 1, K \rrbracket : (p'_i)_{norm}(j) =: \begin{cases} \frac{p'_i(j)}{Z_i} & \text{si } Z_i \neq 0 \\ p'_i(j) & \text{sinon.} \end{cases}$$

Chaque valeur $(p'_i)_{norm}(j)$ non nulle des profils $(p'_i)_{norm}$ en résultant représente ainsi la proportion des connexions entre le point v_i du patch et la région V_j dans l'ensemble des connexions du point v_i .

Dans le cas où l'on applique la PCA sur un groupe de sujets, cela permet aussi d'avoir des profils de connectivité comparables entre les sujets. Pour chaque sujet s des N sujets de la base de sujets considérée, nous noterons ainsi X_{norm}^{ts} la matrice de connectivité associée à ces profils.

2. Analyse en composantes principales de la matrice réduite X'

Intra-sujet Une analyse en composantes principales est ensuite effectuée en utilisant le langage R. Cette analyse projette les profils de connectivité p'_i sur un sous-espace de $(\mathbb{R}^+)^V$ définie par C composantes principales. Le nombre C de ces composantes principales est fixé ici de manière arbitraire dans un premier temps. Nous verrons par la suite des exemples sur les données simulées ainsi que sur les données réelles.

Inter-sujets Dans le cas inter-sujets, une autre étape de **concaténation des profils de connectivité des différents sujets** se rajoute avant l'analyse en composantes principales. Les profils de connectivité normalisés de l'ensemble des sujets sont ainsi concaténés, aboutissant à une matrice X_{norm}^{concat} de taille $(N \times P, K)$. L'analyse en composantes principales est alors effectuée sur cette matrice.

5.3.3 Données Simulées : Cas optimal Template T_1

Intra-sujet Rappelons que le cas optimal (cf. partie Simulation 4.3.2.1 et Fig 5.7) des données simulées a les propriétés suivantes :

- une densité de tracts élevée (10 *tracts/mm²*),
- la distribution locale des connexions d'une parcelle centrée autour du centroïde avec $\sigma = 5 \text{ mm}$ et $\gamma = 0.5$,
- Chaque parcelle a **trois connexions** vers trois parcelles de la surface corticale.

Réduction avant ACP La parcellisation aléatoire utilisée pour la première réduction de dimension avant ACP est une parcellisation de Voronoï centroïdale initialisée avec 200 graines prises aléatoirement sur la surface corticale simulée (cf. Fig. 5.22).

FIG. 5.22 – Parcellisation en 200 régions par Voronoï Centroidal. (cf. partie 3.2.1 pour plus de détails sur la construction de cette parcellisation).



Résultats de l'ACP Une analyse en composantes principales est effectuée d'après la méthode présentée précédemment. Nous trouvons ainsi :

- que la variance des données est expliquée principalement par les quatre premières composantes (cf. Fig 5.23),
- qu'il existe un lien direct entre les vecteurs principaux trouvés et les parcelles connectées au gyrus : ces parcelles correspondent aux maximas et minimas des composantes (cf. Fig. 5.24) ;
- qu'il y a un bon pouvoir de séparation entre les parcelles du gyrus (cf. Fig. 5.25).

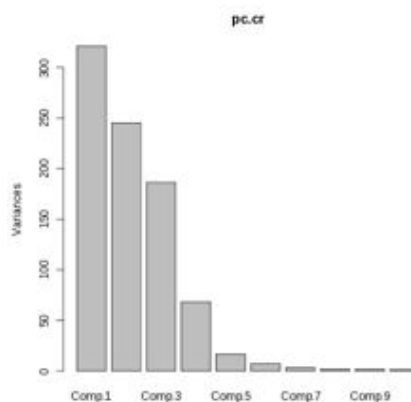


FIG. 5.23 – Cas intra-sujet : ACP du gyrus 1 du sujet 1 de la base B_{T_1} (10 mm). La variance des données de connectivité suivant les différentes composantes de l'ACP. Ici les quatre premières composantes expliquent la majeure partie de la variance.

Inter-sujets Rappelons que les données simulées des 10 sujets de la base B_{T_1} (10) ont les mêmes caractéristiques que le sujet *modèle* T_1 (densité élevée de tracts, distribution locale des connexions d'une parcelle centrée autour du centroïde), avec une variabilité au niveau des parcelles de chaque sujet (les centroïdes de chaque parcelle étant décalés de 10 mm maximum par rapport au sujet *modèle*). Pour plus de détails, veuillez vous reporter à la description des données simulées 4.3.

Les données de connectivité du gyrus 1 de chaque sujet est traité de la façon décrite à la partie 5.3.2. La parcellisation aléatoire utilisée pour réduire les données avec l'ACP est la même que dans le cas intra-sujet précédent.

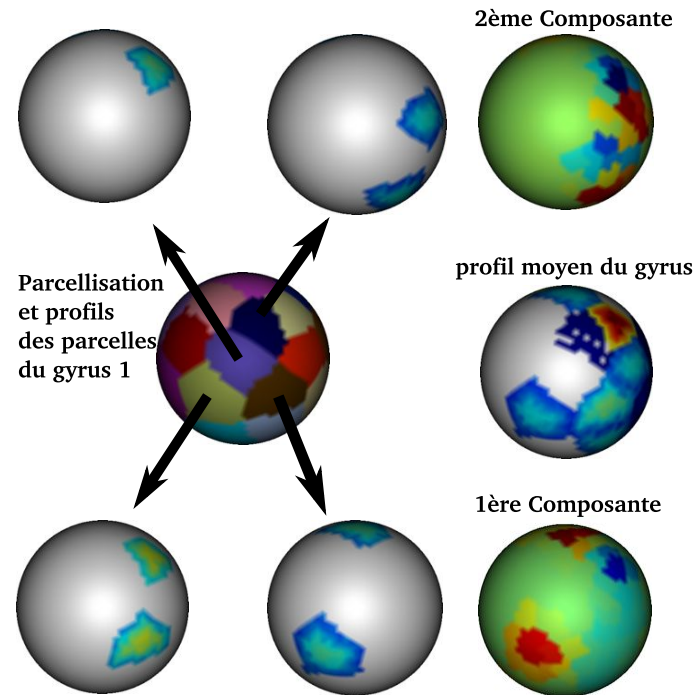


FIG. 5.24 – Cas intra-sujet : ACP du gyrus 1 du sujet 1 de la base B_{T_1} (10 mm). La première composante correspond aux connexions de la parcelle marron, la deuxième à celles de la bleu foncé (termes positifs en rouge). Les termes négatifs (en bleu) correspondent aux connexions des parcelles violette et beige.

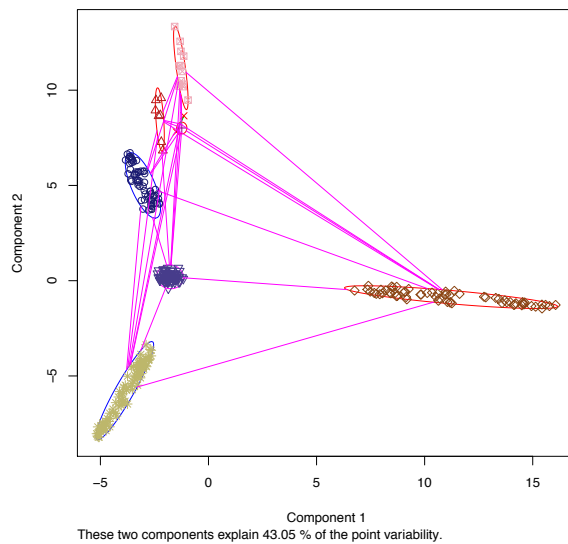
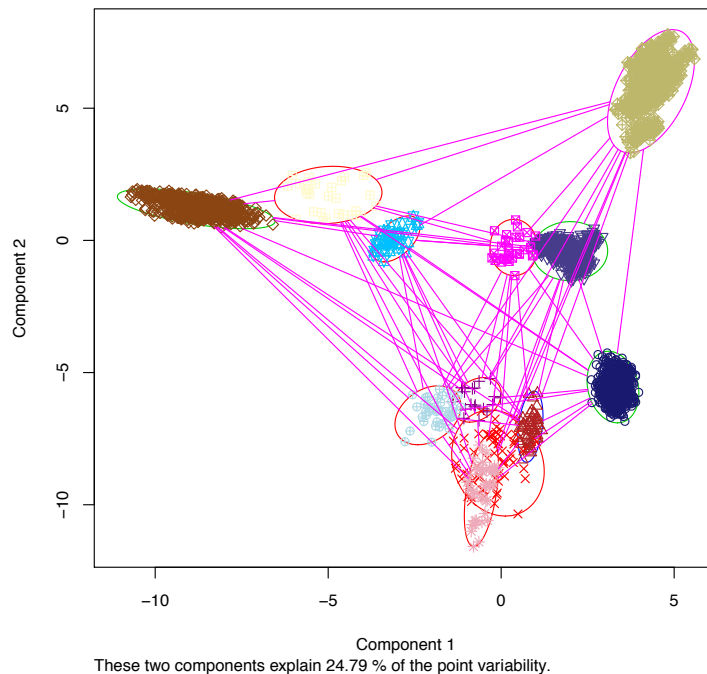


FIG. 5.25 – Cas intra-sujet : ACP du gyrus 1 du sujet 1 de la base B_{T_1} (10 mm). Projection des profils de connectivité des points suivant les deux premières composantes. La couleur des points correspond aux parcelles du gyrus. Les parcelles principales sont en bleu foncé, marron, violet et beige. Les parcelles rose et rouge ne représentent que peu de points (cf. Fig. 5.7(a) pour visualiser leur intersection avec le gyrus 1).

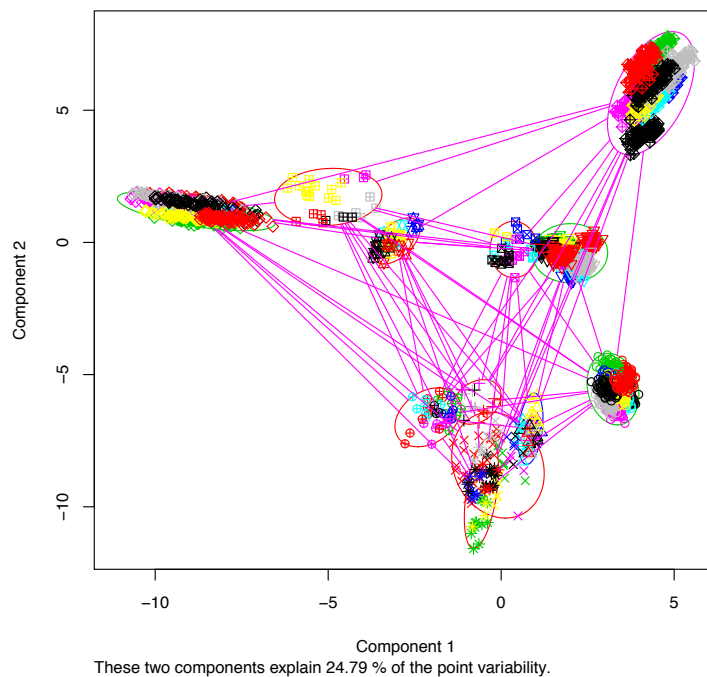
Des constatations similaires au cas intra-sujet peuvent être faites, notamment :

- la variance des données est expliquée principalement par les quatre premières composantes (cf. Fig 5.27),
- les parcelles principales sont déjà bien séparées à partir des deux premières composantes principales (cf. Fig. 5.26(a)),
- C'est en revanche moins le cas pour les parcelles plus petites en périphérie du gyrus, qui ne sont pas présentes chez tous les sujets,

- Pour les parcelles principales, les profils projetés suivant les deux premières composantes sont regroupés par sujets (cf. Fig. 5.26(b)).



(a)



(b) La couleur des points correspond aux différents sujets.

FIG. 5.26 – Cas inter-sujets : ACP du gyrus 1 de la base B_{T_1} (10 mm). Projection des profils de connectivité des points suivant les 2 premières composantes.

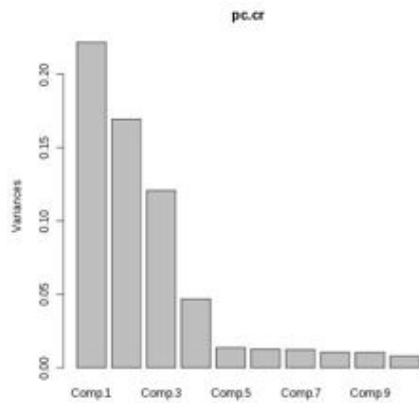


FIG. 5.27 – Cas Inter-sujets : ACP du gyrus 1 de la base B_{T_1} (10 mm). La variance des données de connectivité suivant les différentes composantes de l'ACP. Les 4 premières composantes expliquent la majeure partie de la variance.

5.3.4 Données simulées : Cas extrême

Rappels Le sujet *modèle* du cas extrême des données simulées a pour caractéristiques :

- une connexion seulement par parcelle,
- une densité de tracts faible (0.3 tracts/mm^2),
- la distribution locale des connexions d'une parcelles est moins gaussienne (autour du centroïde) avec $\sigma = 5 \text{ mm}$ et $\gamma = 0.9$.

A partir de ce *modèle*, une base de sujets avec une variabilité assez importante ($d_{\max} = 20 \text{ mm}$) est construite : $B_{T_{0.3}, \sigma=5, \gamma=0.9}(20 \text{ mm})$.

Résultats L'analyse en composantes principales ne donne pas des résultats aussi satisfaisants que précédemment, aussi bien dans le cas intra-sujet que inter-sujets :

- la variance des données est expliquée par un nombre plus important de composantes (cf. Fig. 5.28(a) et 5.28(b)),
- la séparation entre les parcelles est moins évidente (cf. Fig. 5.29 et 5.30).

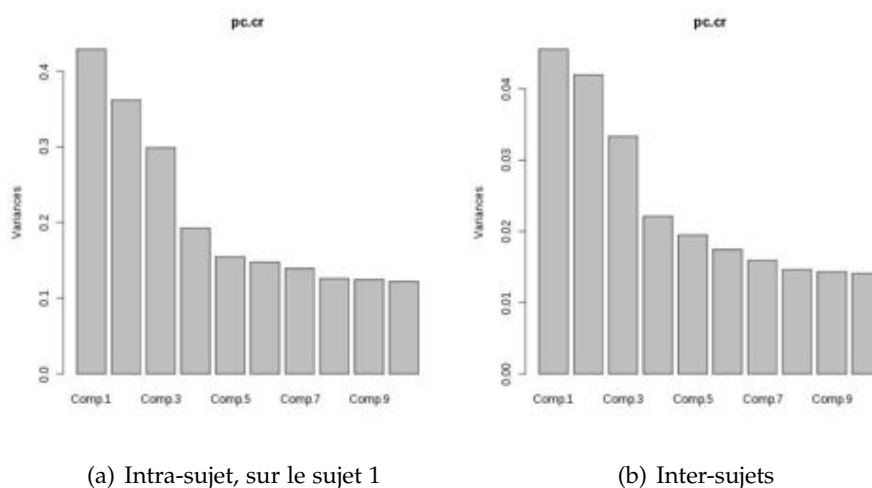


FIG. 5.28 – Variance expliquée par les composantes de l'ACP du gyrus 1 de la base extrême. La variance des données de connectivité est expliquée par un nombre plus important de composantes que dans le cas des données simulées optimales.

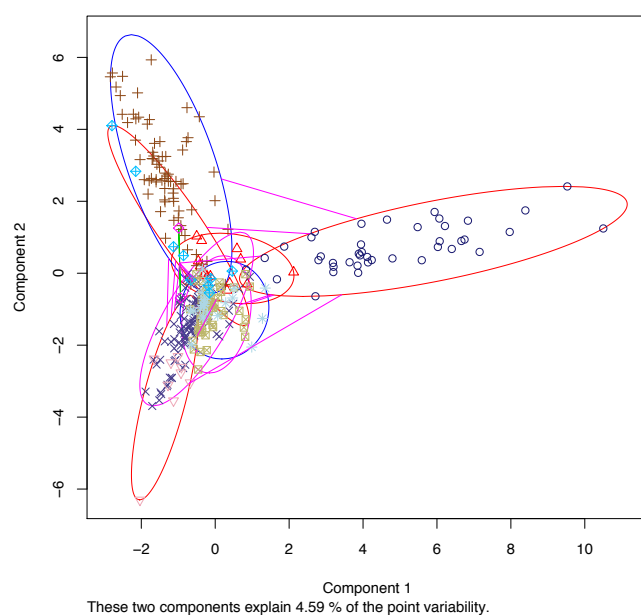


FIG. 5.29 – Cas Intra-sujet : ACP du gyrus 1 du sujet 1 de la base extrême. Projection des profils de connectivité des points suivant les deux premières composantes.

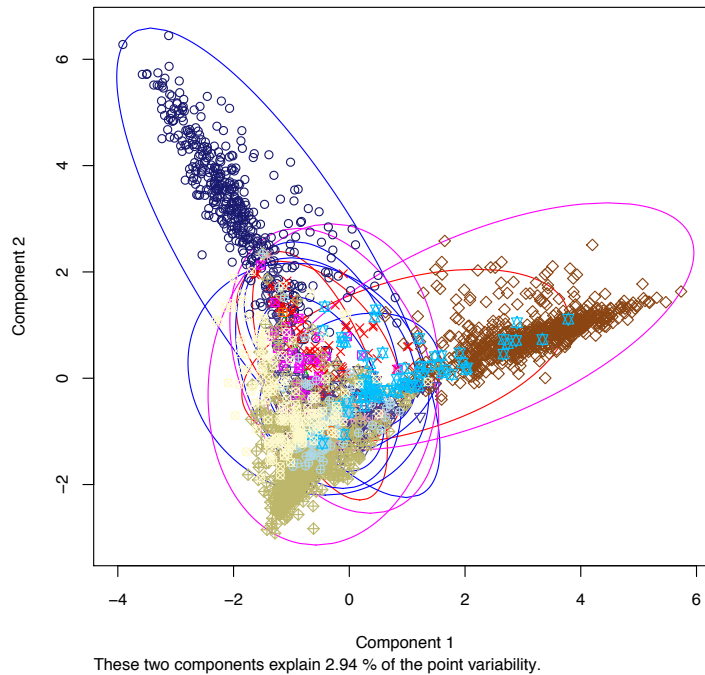


FIG. 5.30 – Cas Inter-sujets : ACP du gyrus 1 de la base extrême. Projection des profils de connectivité des points suivant les deux premières composantes.

5.3.5 Essais sur les données réelles

Nous présentons ici l'exemple d'une analyse en composantes principales des données de connectivité du gyrus post-central droit.

Gyrus post-central droit

Cas Intra-sujet : Exemple du sujet 1

Données d'entrée Les données utilisées sont celles présentées à la partie 2.4.2). Dans ce contexte :

- le patch surfacique correspondant au gyrus post-central droit possède environ $P = 2200$ points ,
- la parcellisation utilisée pour réduire les données est découpée en environ $K = 2400$ régions par hémisphère,
- la matrice de connectivité est de plus normalisée conformément à la description précédente (cf. 5.3.2).

Pour simplifier le problème l'ACP est effectuée à partir de la matrice de connectivité vers l'hémisphère droit (avec suppression des régions appartenant au gyrus post-central). Les données d'entrée sont donc sous la forme de la matrice suivante :

$$\mathbf{X}' = \begin{pmatrix} V_1 & \dots & V'_K \\ \dots & p'_1 & \dots \\ & \vdots & \\ \dots & p'_P & \dots \end{pmatrix} \begin{matrix} v_1 \\ v_P \end{matrix}$$

avec $\{\mathcal{V}_k\}_{k=1}^{K'}$ l'ensemble des régions de la segmentation de l'hémisphère droit qui sont connectées au gyrus ($K' < P < K$).

Résultats Après l'analyse en composantes principales, nous constatons que :

- en comparaison aux données simulées, la variance des données est expliquée d'une façon intermédiaire par rapport aux deux bases (optimale et extrême) ; en effet, cette variance est expliquée par un plus grand nombre de composantes que pour la base simulée optimale, mais la variance expliquée décroît plus rapidement en fonction du numéro de la composante par rapport à la base simulée extrême (cf. Fig. 5.31) ;
- la visualisation sur la surface corticale des trois premières composantes principales est en lien avec les connexions du gyrus comme dans le cas des données simulées même si cela est moins clair (cf. Fig. 5.32 et 5.33) ;
- les scores des points du gyrus suivant les premières composantes permettent l'identification de différentes régions spatiales sur le gyrus (cf. Fig. 5.34). Les composantes ont un pouvoir de séparation des points.

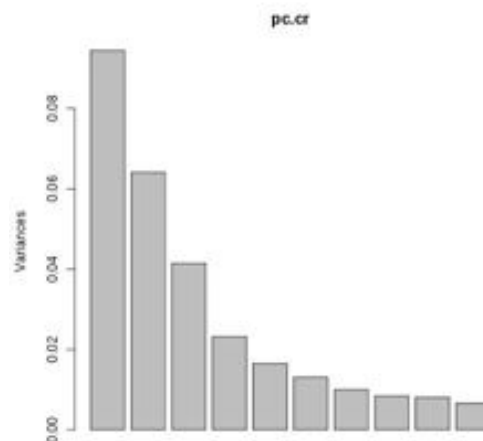


FIG. 5.31 – ACP du gyrus post-central droit du sujet 1 à partir des données de connectivité anatomique de l'hémisphère droit. Variance représentée par les 10 premières composantes.

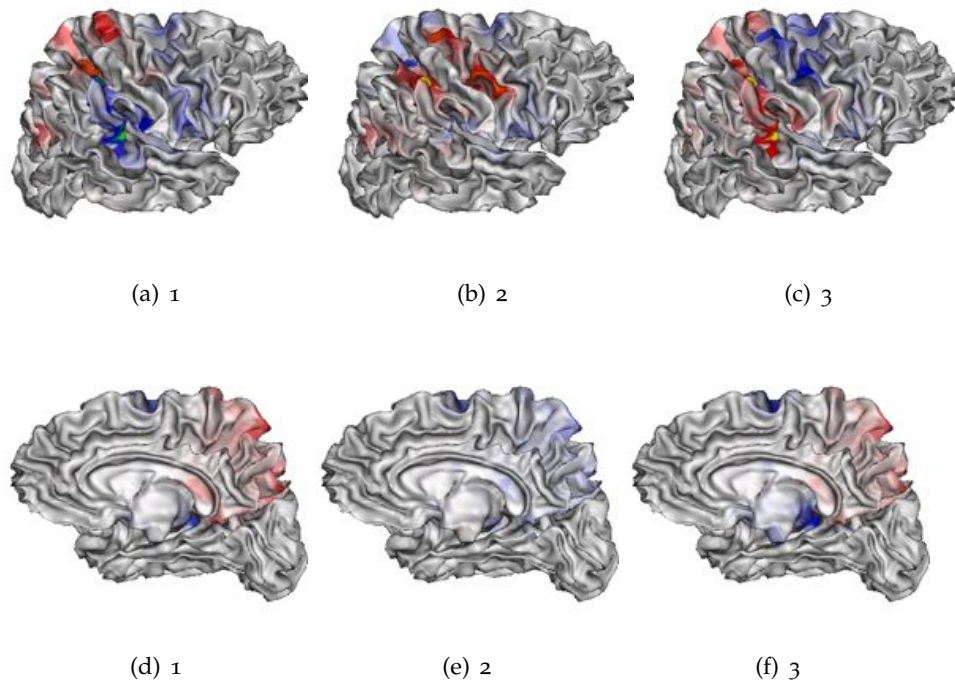


FIG. 5.32 – ACP du gyrus post-central droit du sujet 1 (sur les données de connectivité de l'hémisphère droit). Visualisation des 3 premières composantes principales en vue de côté (5.32(a), 5.32(b), 5.32(c)) et vue médiale (5.32(d), 5.32(e), 5.32(f)). Les couleurs représentent le poids de chacune des composantes suivant les points de la surface corticale (caractérisant l'espace non réduit des profils de connectivité complets). Nous pouvons noter qu'il existe un lien entre les composantes obtenues et les régions connectées au gyrus (Fig. 5.33).

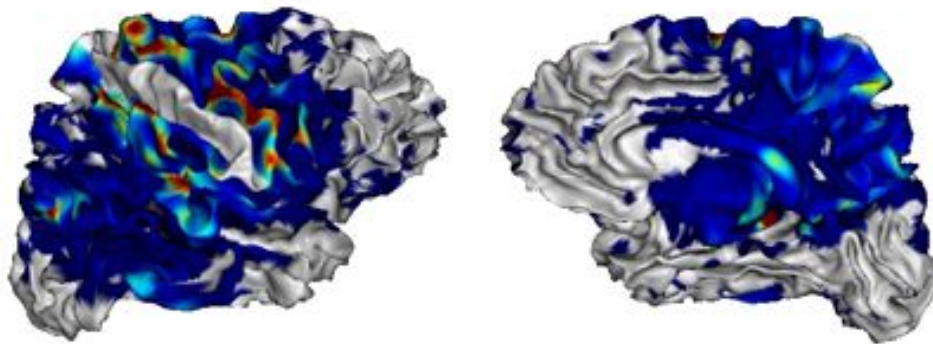


FIG. 5.33 – Profil de connectivité du gyrus post-central droit du premier sujet de la base NMR. Cette visualisation surfacique du profil du gyrus permet d'identifier un ensemble de régions d'intérêt correspondant à l'arrivée de chaque faisceau. Ces régions sont en lien avec les composantes principales de l'ACP effectuée sur les données de connectivité du gyrus (Fig. 5.32). Rappelons que nous avons présenté une façon simple de les identifier automatiquement dans la partie 5.2.

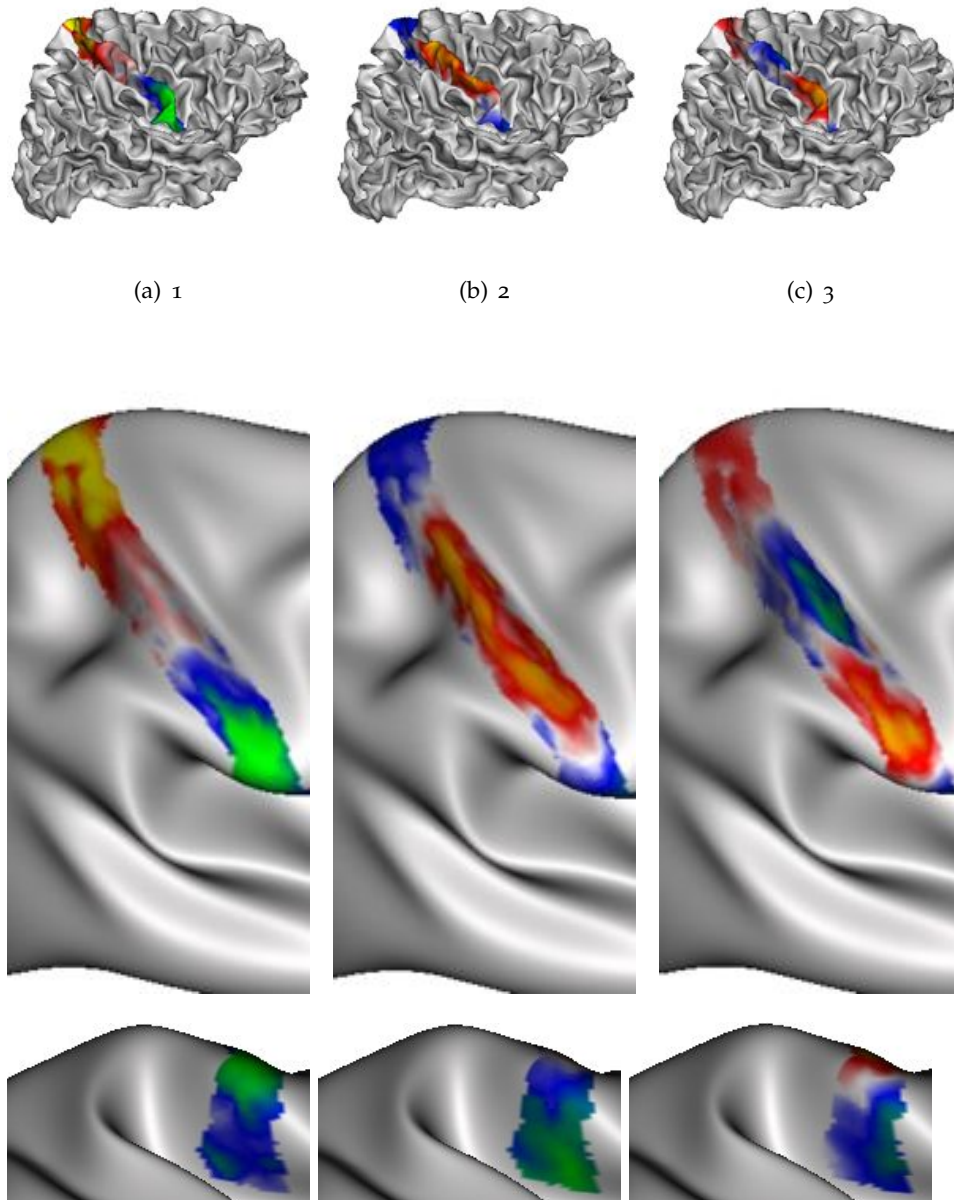


FIG. 5.34 – ACP du gyrus post-central droit du sujet 1 à partir des données de connectivité anatomique de l'hémisphère droit. Visualisation des scores des points du gyrus suivant les 3 premières composantes principales (5.34(a), 5.34(b), 5.34(c)). La visualisation est effectuée sur le maillage natif du sujet ou sur le maillage gonflé avec un zoom.

5.4 DIFFÉRENTES MESURES DE DISTANCE ENTRE PROFILS

Dans cette section nous allons aborder différentes mesures de distance ou de similarités possibles entre deux profils de connectivité. A l'aide des données simulées, nous évaluerons alors leur pertinence, et leur conséquence sur le clustering intra-sujet (partie 6) et inter-sujets (partie 7.1).

Notations, rappel En considérant l'ensemble $\mathcal{V} = \{v_i, \dots, v_V\}$ des V points d'un maillage du cortex, nous avons précédemment noté $p_i \in$

$(\mathbb{N}^+)^V$) le profil de connectivité d'un point v_i de \mathcal{V} vers l'ensemble des points de \mathcal{V} . Dans la suite, nous considérerons deux profils de connectivité appelés $x = [x_1, \dots, x_V]$ et $y = [y_1, \dots, y_V]$ correspondant à deux profils p_i et p_j .

5.4.1 Définitions

5.4.1.1 La corrélation

Dans la plupart des études récentes parcellisant le cortex en fonction de la connectivité anatomique, une **matrice de cross-correlation** est calculée à partir de la matrice de connectivité obtenue par tractographie probabiliste (cf. chapitre introductif 1.2.3).

La corrélation (Pearson's Correlation) entre deux variables aléatoires X et Y d'espérances μ_X et μ_Y et d'écarts type σ_X et σ_Y respectivement, est définie comme :

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (5.2)$$

Si l'on dispose d'un échantillon de N mesures $x = (x_i)_{i=1..N}$ et $y = (y_i)_{i=1..N}$ de X et Y , une estimation de la corrélation de Pearson entre X et Y est donnée par le coefficient de corrélation de l'échantillon défini comme :

$$r = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{x_i - \bar{x}}{s_X} \right) \left(\frac{y_i - \bar{y}}{s_Y} \right) \quad (5.3)$$

où \bar{x} et s_X , \bar{y} et s_Y sont les moyennes et écarts type des échantillons x et y respectivement.

5.4.1.2 Distance euclidienne

Etant donnés deux profils de connectivité X et Y , la distance euclidienne entre ces deux vecteurs est définie par :

$$d(X,Y) = \sqrt{\sum_{i=1}^V (x_i - y_i)^2} \quad (5.4)$$

5.4.1.3 Distance de Hausdorff

La distance de Hausdorff est une mesure de distance entre deux sous-ensembles de points associés à un espace métrique.

Soit X et Y deux sous-ensembles non vides d'un espace métrique (M, d) . La distance de Hausdorff $d_H(X, Y)$ est définie par :

$$d_H(X, Y) = \max\left\{\sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y)\right\} \quad (5.5)$$

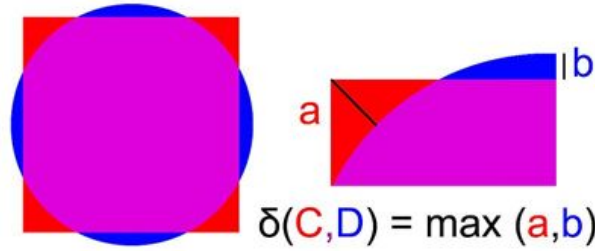


FIG. 5.35 – Principe de la distance de Hausdorff entre deux ensembles C (carré rouge) et D (disque bleu).

L'intérêt de cette distance est qu'elle tient compte du côté spatial. Comme le montre la figure 5.36, l'**aspect spatial** des données et la distance spatiale correspondante est prise en compte lors du calcul de la distance de Hausdorff alors que cela n'est pas le cas pour la distance euclidienne. Cette figure présente trois exemples de profils. Ainsi le profil 1 a deux connexions : une rouge et une bleue (les couleurs permettent d'apparier les connexions deux à deux les plus proches sur les trois profils. Les distances entre ces couples de connexions seront utilisées dans le calcul de la distance de Hausdorff.). Les deux autres profils ont deux connexions décalées par rapport au premier profil. La distance euclidienne entre chacun des profils p_1 et p_2 avec le premier profil est également élevée : $d_{euclid}(p_1, p_2) = d_{euclid}(p_1, p_3)$. En revanche, la distance de Hausdorff est différente dans les deux cas, avec $d_H(p_1, p_2) < d_H(p_1, p_3)$. Ceci correspond bien à l'appréciation visuelle que nous faisons lorsque nous comparons les deux profils : le pattern spatial formé par le profil 2 est plus proche du profil 1 que ne l'est celui du profil 3.

Nous avons vu au 5 une méthode de réduction des profils de connectivité permettant de s'affranchir d'une partie de la variabilité inter-individuelle. Une autre idée serait de définir une autre distance quantifiant la variabilité entre les profils, n'étant pas nulle lorsque les profils ne se superposent pas, mais fonction du décalage spatial de leur pattern surfacique. La distance de Hausdorff semble répondre à ces critères.

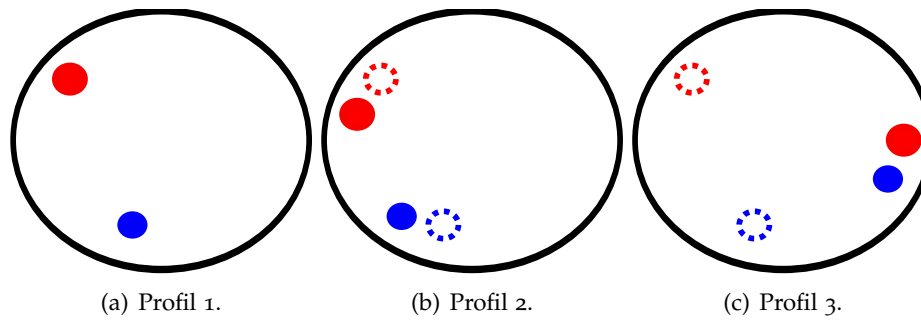


FIG. 5.36 – Illustration de la différence entre les distances euclidienne et de Hausdorff avec des exemples de profils de connectivité : Trois exemples de profils. Les deux autres profils ont deux connexions décalées par rapport au premier profil (cf. trace en pointillés du profil 1). La distance euclidienne entre chacun des profils p_1 et p_2 avec le premier profil est également élevée : $d_{euclid}(p_1, p_2) = d_{euclid}(p_1, p_3)$. Dans le cas de la distance de Hausdorff, cette distance est différente dans les deux cas, avec $d_H(p_1, p_2) < d_H(p_1, p_3)$. Ceci correspond bien à l'appréciation visuelle que nous faisons des patterns formés lorsque nous comparons les deux profils.

5.4.2 Données Simulées

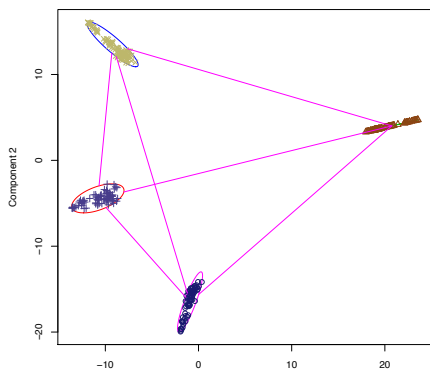
Cas optimal

Nous avons ainsi calculé les distances entre les profils de connectivité du gyrus 1 du cas optimal des données simulées pour les trois types de distances définis précédemment. La figure 5.37 montre les résultats d'une représentation bidimensionnelle des points du gyrus construite à partir des différentes matrices de distances associées.

La prise en compte du côté spatial des données est bien illustrée dans ce cas. En effet, les parcelles bleu foncé et beige ont des connexions différentes (ne se chevauchant pas spatialement) :

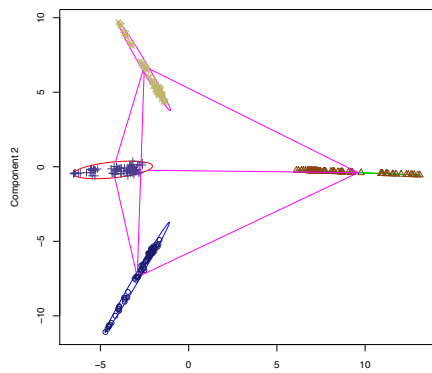
- elles sont donc bien séparées par les deux premières mesures,
- la distance de Hausdorff en revanche les sépare moins et ceci est cohérent car ces parcelles présentent des connexions qui sont deux à deux proches spatialement de manière géodésique à la surface corticale (cf. Visualisation des connexions à la figure 5.24).

FIG. 5.37 – Cas optimal : Représentation bidimensionnelle des points du sujet modèle (faite par scaling multidimensionnel) à partir de 3 types de distances entre les profils de connectivité associés. Une couleur différente est associée aux points de chaque parcelle.



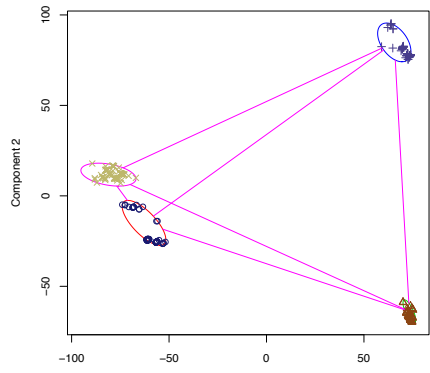
These two components explain 85.67 % of the point variability.

(a) Cross Correlation, avec PCA



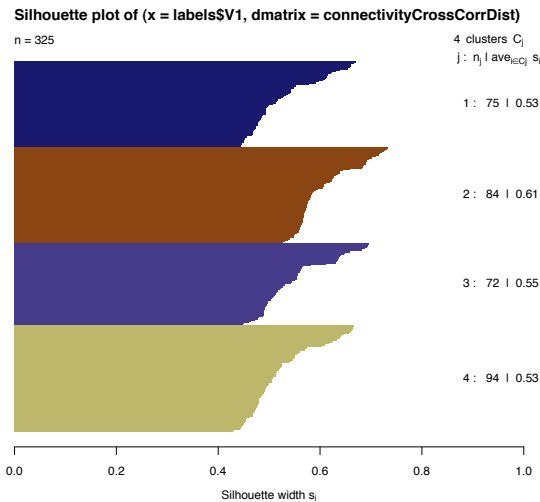
These two components explain 45.27 % of the point variability.

(b) Distance euclidienne, avec Multi.Dim.scaling

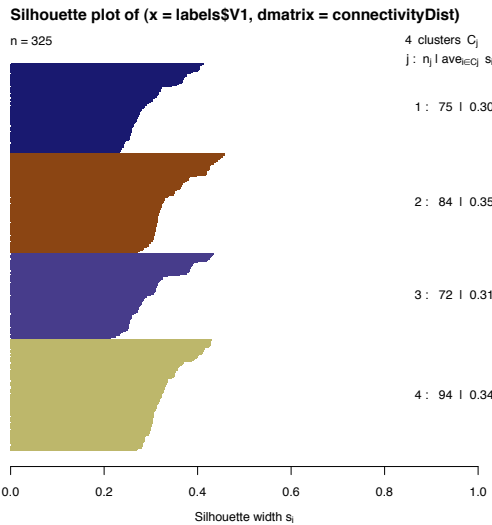


These two components explain 27.66 % of the point variability.

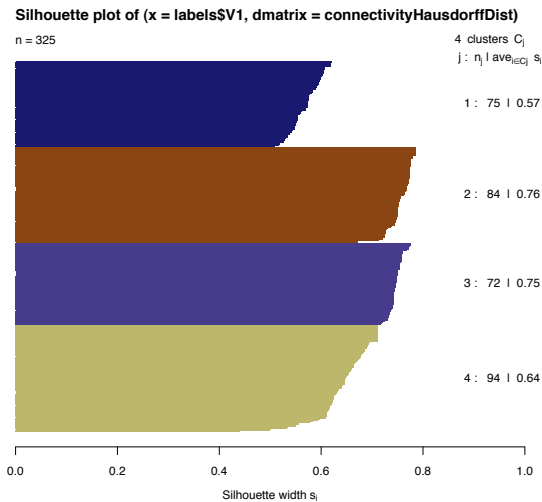
(c) Distance de Hausdorff, avec Multi.Dim.scaling



(d) Cross Correlation (avg=0.55)



(e) Distance euclidienne (avg=0.33)



(f) Distance de Hausdorff (avg=0.68)

5.4.2.1 Cas extrême

Rappels Le sujet *modèle* du cas extrême des données simulées a pour caractéristiques (cf. partie 4.3.2.2) :

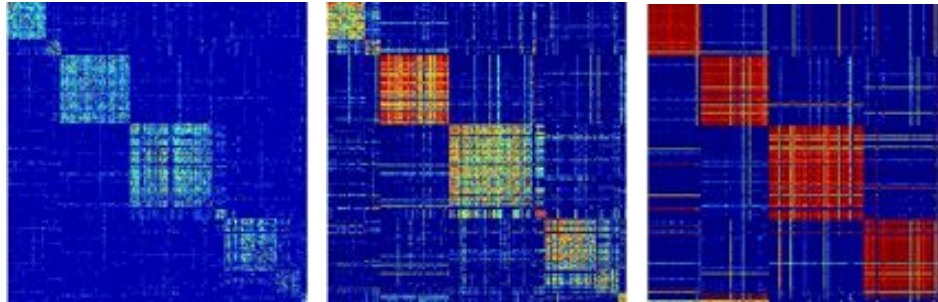
- une connexion seulement par parcelle,
- une densité de tracts faible (0.3 tracts/mm^2),
- la distribution locale des connexions d'une parcelles est moins gaussienne (autour du centroïde) avec $\sigma = 5 \text{ mm}$ et $\gamma = 0.9$.

A partir de ce *modèle*, une base de sujets avec une variabilité assez importante ($d_{\max} = 20 \text{ mm}$) est construite : B_{extreme} .

Effet de la réduction de dimension des profils La figure 5.38 montre la matrice de cross-corrélation des profils de connectivité des points du gyrus dans trois cas :

1. avec les profils réduits à partir d'une parcellisation de Voronoï aléatoire du cortex (cf. Fig. 5.38(g) et partie 5.3.2 pour la réduction de dimension),
2. avec les profils réduits à partir de l'identification automatique des régions d'intérêt connectées au gyrus (cf. Partie 5.2 pour la méthode de réduction) :
 - a. au niveau intra-sujet (cf. Fig. 5.38(h),
 - b. au niveau du groupe de sujet (cf. Fig. 5.38(i)).

Nous constatons que la réduction de dimension par identification de régions d'intérêt renforce les corrélation entre les points du gyrus au sein des parcelles. Avec un effet plus important avec la réduction au niveau inter-sujets.



(g) Profils réduits (parcellisation aléatoire 200 parcelles) (h) Profils réduits en intra-sujet (i) Profils réduits en inter-sujets

FIG. 5.38 – Cas extrême B_{extreme} . Matrice de cross-corrélation entre les profils de connectivité pour 3 cas : 5.38(g) Réduction à partir d'une parcellisation aléatoire du maillage du cortex en 200 parcelles, 5.38(h) Réduction intra-sujet décrite précédemment, 5.38(i) Réduction inter-sujets décrite précédemment.

5.5 CONCLUSION

Dans la littérature, pour parcelliser une région du cortex, deux types de profils sont principalement utilisés, associés à deux approches : l'approche aveugle où l'on prend en compte les profils complets, et l'approche avec a priori où les profils sont regroupés en fonction de régions d'intérêt identifiées à partir d'une segmentation a priori du cortex ou manuellement avec des a priori sur les connexions.

Dans cette section, nous proposons de nouvelles manières de résumer l'information contenue dans les profils de connectivité d'un patch du cortex. Ceci est fait à partir des données de connectivité elle-mêmes, et sans l'utilisation d'une parcellisation a priori du cortex. Nous avons à chaque fois étudié ces approches sur des données simulées et sur des données réelles.

Nous avons ainsi présenté une méthode permettant l'**identification automatique des régions d'intérêt connectées au patch, pour un sujet ou pour un groupe de sujets**. Cette méthode permet : d'isoler les différents tracts connectés au patch et de pouvoir les visualiser, et peut ensuite servir à réduire les profils de connectivité en vue d'effectuer un clustering sur ces profils. Ce type de réduction augmente la similarité entre les profils des points au sein de chaque parcelle dans le cas d'un exemple sur les données simulées. Ceci sera utile pour une parcellisation en fonction de ces profils. Dans le cas des données réelles, grâce à un outil de visualisation développé pendant cette thèse, nous constatons, dans le cas d'un patch englobant le SMA et le pré-SMA, que cette région comporte des forces de connexion différentes suivant les ROIs identifiées précédemment. Ceci sera intéressant à étudier dans le cas d'un clustering de cette zone par la suite.

Nous avons aussi étudié le comportement d'une **analyse en composantes principales sur les données de connectivité**. Dans le cas de données simulées comme dans le cas des données réelles (analyse sur le gyrus post-central), les parcelles connectées au patch sont représentées par les minimas et les maximas des composantes principales. La variance est de plus principalement expliquée par les quatre premières composantes dans le cas du gyrus post-central. Enfin, les scores des points du patch suivant les premières composantes varient spatialement, en formant des zones homogènes, et cohérentes spatialement. Dans la suite du manuscrit, nous essaierons de les identifier automatiquement par clustering.

Profils et parcellisation

Comme l'illustre la figure 5.39, parcelliser le cortex à partir des profils de connectivité complets (approche aveugle) ou à partir de régions d'intérêt bien identifiées peut conduire à des résultats très différents. En effet sur cet exemple, une parcellisation basée sur le **profil de connectivité avec tout le cerveau** rassemble les voxels qui ont des coordonnées rétinotopiques similaires (couleurs claires avec couleurs claires, couleurs foncées avec couleurs foncées). En revanche, si les trois aires rétinotopiques appartiennent à différentes parcelles d'une parcellisation à large échelle, une

classification basée sur le **profil de connectivité vers chaque région** de cette parcellisation va séparer le système en trois aires distinctes.

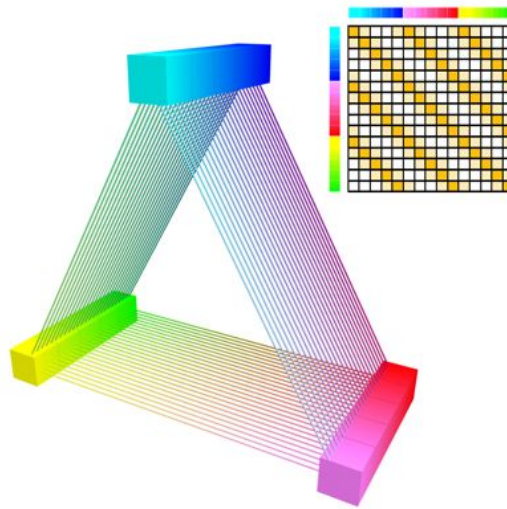


FIG. 5.39 – Réseau simulé de trois aires inspirées du réseau rétinotopique et leur matrice de connectivité. (source : Perrin (2006)).

Il est ainsi intéressant d'étudier le comportement du clustering en fonction des profils de connectivité utilisés (associés à la méthode de réduction de dimension appliquée).

Comme le montre la figure 5.40, en dehors de la définition des profils de connectivité, il existe de multiples démarches pour parcelliser le cortex à partir des données de connectivité anatomique, ainsi que différentes méthodes pour sa réalisation.

Nous allons dans un premier temps présenter une démarche possible avec un choix de méthodes donné (Encadrés en rouge sur la figure 5.40). Cette démarche a abouti à un article de conférence (Roca et al. 2009a) et sera abordée dans le chapitre 6 suivant. Elle propose une parcellisation du cortex entier, en intra-sujet, avec une démarche itérative, en réduisant les profils à partir de l'identification automatique de régions d'intérêt définie précédemment (cf. partie 5.2.1).

Dans un deuxième temps nous proposerons une méthode parcellisant une région donnée au niveau d'un groupe de sujets, utilisant une réduction de dimension des profils de connectivité grâce à l'identification automatique de régions d'intérêt définies au niveau du groupe (cf. partie 5.2.2 et Encadrés en jaune sur la figure 5.40). Cette approche a abouti à un article de conférence (Roca et al. 2010a) et sera abordée dans la partie 7.1.

Dans un troisième temps (partie 7.2), nous aborderons les autres démarches testées durant ma thèse, démarches se différenciant par la technique de réduction de dimension utilisée. Nous comparerons les résultats obtenus avec les différentes méthodes sur les données simulées ainsi que sur les données réelles.

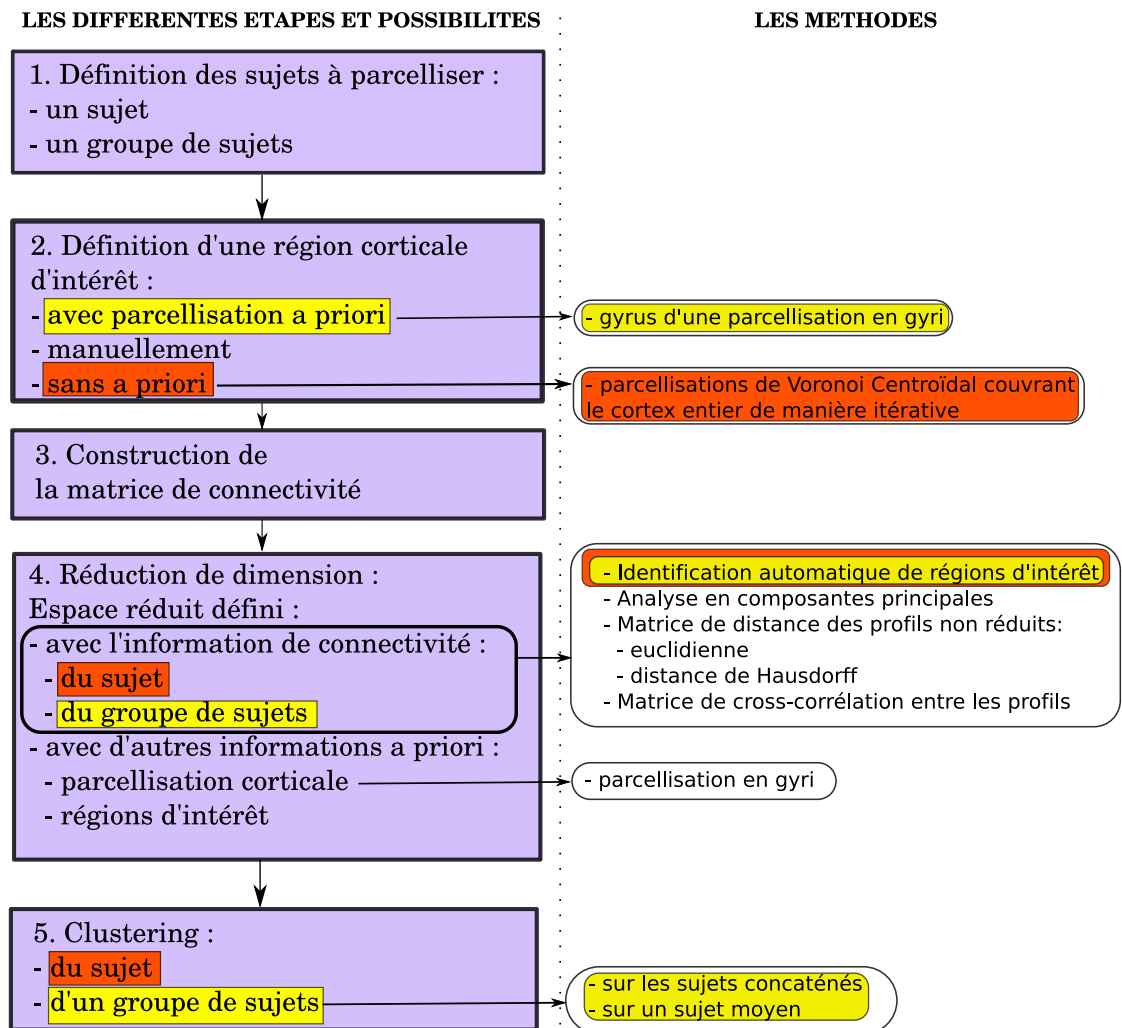


FIG. 5.40 – Parcellisation du cortex en fonction des données de connectivité anatomique. Exemple de différentes approches envisageables avec les méthodes testées. L'approche constituée des encadrés oranges sera abordée dans le chapitre 6, celle des encadrés jaunes dans le chapitre 7.1.

PARCELLISATION ITÉRATIVE INTRA-SUJET DU CORTEX ENTIER

L'OBJECTIF de ce travail est d'effectuer une parcellisation du **cortex entier** en fonction des données de connectivité anatomique pour **chaque sujet indépendamment** et ceci **sans utiliser une segmentation a priori** pour réduire les profils de connectivité.

Nous proposons ici un **algorithme itératif**, avec une **réduction de dimension de la matrice de connectivité** à deux niveaux : (i) avec une **division du domaine à parcelliser** à chaque itération, (ii) en réduisant les données de connectivité de ce sous-domaine par l'**identification automatique de régions d'intérêt** (cf. méthode décrite dans la partie 5.2 précédente). Ceci est réalisé sans utiliser de segmentation a priori du cortex et aboutit à une parcellisation du cortex entier.

La démarche est très ambitieuse et nous avons fait des hypothèses pour la simplifier, notamment par rapport à la longueur des tracts pris en compte. Les régions du cortex ne présentant pas ou peu de connexions longues distance ne sont pas parcellisées. Cette étude est de plus réalisée sur un nombre limité de sujets (4) et avec une tractographie déterministe, ce qui limite ainsi la quantité de données à traiter.

Ce travail a abouti à un article de conférence (Roca et al. 2009a).

SOMMAIRE

7.1	PARCELLISATION AVEC L'IDENTIFICATION DE RÉGIONS D'INTÉRÊT	147
7.1.1	La méthode	147
7.1.2	Essais sur les données simulées	151
7.1.2.1	Cas optimal	151
7.1.2.2	Cas extrême	153
7.1.3	Essais sur les données réelles	154
7.1.4	Discussion	157
7.2	COMPARAISON DE DIFFÉRENTES MÉTHODES	158
7.2.1	Récapitulatif des approches	158
7.2.2	Procédure	158
7.2.3	Résultats sur données simulées	159
7.2.4	Données réelles : région englobant le SMA et le pré-SMA	164
7.2.4.1	Méthode M_{ROIs}^{inter}	164
7.2.4.2	Méthode M_{PCA}^{intra}	168
7.2.4.3	Discussion	170
7.3	CONCLUSION	171

6.1 DONNÉES ET HYPOTHÈSES

Considérons un sujet, avec comme données la surface corticale extraite des données anatomiques comme décrit précédemment, ainsi que les tracts et les matrices de connectivité associées.

Données d'entrée de la méthode La méthode prend en entrée :

- la surface corticale $S = (\mathcal{V}, \mathcal{E})$, où $V = \text{card}(\mathcal{V})$
- la matrice de connectivité $C \in \mathcal{M}_{(\mathcal{V}_{init}, \mathcal{V})}(\mathbb{R})$ associée à S_{init} , qui définit par point $v_i \in \mathcal{V}_{init}$ un profil de connectivité quantifiant sa force de connexion vers chaque point $v_j \in \mathcal{V}$.

Hypothèses Pour simplifier le problème, nous avons été amenés à faire des hypothèses :

- Nous supposons que les **tracts longs sont plus pertinents** pour parcelliser que les tracts courts en faisant l'hypothèse que les tracts longs sont plus reproductibles à travers les individus, et qu'ils contiennent plus d'informations pertinentes pour le clustering (la répartition des tracts courts dépendant principalement de la géométrie du maillage),
- Nous posons un a priori sur la **moyenne de la surface d'une aire** (parcelle) à trouver, moyenne notée $moyenne(surf_{parcelle})$. Par exemple pour le cortex, si l'on cherche environ 150 aires par hémisphère, chaque aire corticale individuelle devrait occuper en moyenne 6 cm^2 avec des aires beaucoup plus grandes et d'autres beaucoup plus petites (grand écart type) Van Essen et Dierker (2007);
- On fixe aussi la **surface minimale d'une aire**, notée $minSurf_{parcelle}$, qui sera prise égale à 5 mm^2 .

Définition de la surface corticale d'intérêt à parcelliser La surface initiale d'intérêt à parcelliser ($S_{init} \subset S$) peut être définie de différentes façons :

- soit on ne prend en compte aucune contrainte et on parcellise l'ensemble de la surface : $S_{init} = S$,
- soit un nombre minimal de tracts intersectant un point de la surface est imposé et dans ce cas tous les points ne satisfaisant pas ce critère sont enlevés de la surface à parcelliser S ,
- soit une région d'intérêt définie par une segmentation a priori (anatomique via les gyri par exemple) est utilisée.

Nous verrons par la suite que nous avons choisi la deuxième option dans le cas des données réelles.

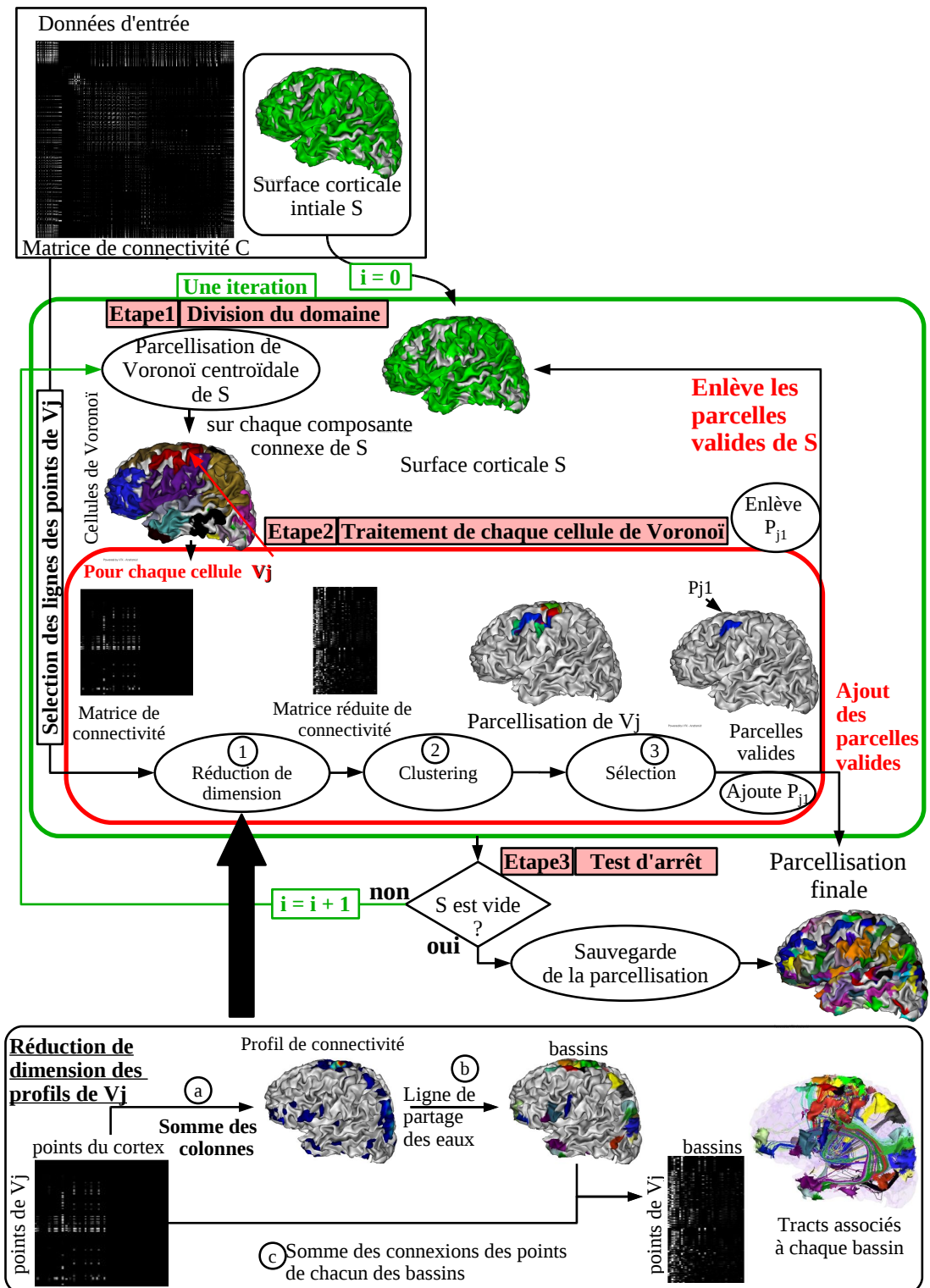


FIG. 6.1 – Organigramme de la méthode. Dans le traitement d'une cellule, la première étape de réduction de dimension peut être soit la technique de réduction des profils de connectivité à partir de l'identification de régions d'intérêt (cf. 5.2), soit une autre technique de réduction de dimension, par exemple une ACP.

6.2 MÉTHODE

Principe

Une itération L'algorithme est itératif. Chaque itération consiste en deux étapes :

1. Division du domaine à parcelliser S_i en *cellules de Voronoï* (premier niveau de réduction de dimension) en appliquant une parcellisation de Voronoï centroïdale (cf. 3.2.1) ;
2. Parcellisation de chaque cellule C_j en parcelles de profils de connectivité homogènes.

Chaque parcelle n'intersectant pas la frontière de la cellule est sauvegardée dans la parcellisation finale. Avant le clustering de chaque *cellule*, les profils de connectivité sont de plus réduits en des profils plus courts vers un ensemble de régions d'intérêt identifiées à partir des données de connectivité comme expliqué précédemment 5.2 (deuxième niveau de réduction de dimension). Une fois toutes les cellules traitées, on met à jour le domaine restant à parcelliser à l'itération suivante en fonction des parcelles identifiées.

Critère d'arrêt On réitère le processus sur le domaine \mathcal{V}_{i+1} jusqu'à ce que ce domaine soit vide.

Le processus itératif de parcellisation de la surface d'intérêt initiale S_{init} aboutira à une parcellisation finale, $P_{parcellisation} = (\mathcal{V}_{parcellisation}, \mathcal{L}_{parcellisation})$

où : $\begin{cases} \mathcal{V}_{parcellisation} \subset \mathcal{V}_{init} \text{ représente l'ensemble des points } v_i \text{ appartenant à une parcelle} \\ \mathcal{L}_{parcellisation} \text{ représente les labels des parcelles associées aux } v_i. \end{cases}$

L'algorithme 3 et la figure 6.1 donnent un aperçu du principe de la méthode que l'on va détailler étape par étape, pour une itération i .

6.2.1 Division du domaine

En pratique, le domaine S_i à parcelliser au début de l'itération i peut comporter plusieurs composantes connexes. Dans un premier temps, ces composantes sont d'abord identifiées. Chacune de ces composantes est ensuite subdivisée en *cellules*. L'ensemble des cellules de toutes les composantes connexes forme une partition $\cup_{j=1..J} C_j$ de S_i .

Chaque composante est découpée ainsi en *cellules* uniformément réparties en appliquant un algorithme de Voronoï centroïdal (cf. 3.2.1 pour une description de l'algorithme utilisé).

Si l'on fait une hypothèse sur la surface moyenne d'une parcelle *moyenne(surf_{aire})*, on peut déterminer le nombre de *cellules* par composante de manière à ce que chaque *cellule* contienne au moins 10 aires corticales. En notant *surf_{composante}* la surface de la composante connexe, le nombre de cellules J sera donc déterminé de la façon suivante :

$$J = \text{partie Entiere}(\text{surf}_{composante} / \text{moyenne}(\text{surf}_{aire}))$$

Algorithme 3: Parcellisation itérative du cortex entier d'un sujet

```

 $i = 0$ 
 $\mathcal{V}_i = \mathcal{V}_{init}$ 
 $\mathcal{V}_{parcellisation} = \emptyset$ 
 $\mathcal{L}_{parcellisation} = \emptyset$ 
while  $\mathcal{V}_i \neq \emptyset$  and  $i \leq 100$  do
  {ETAPE 1 : DIVISION DU DOMAINE EN CELLULES DE VORONOI}
  Division du domaine à parcelliser  $S_i$  et des points associés  $\mathcal{V}_i$  en
  cellules de Voronoï  $C_j$  avec :
     $S_i = \cup_{j=1..J} C_j$  avec  $\cap_{j=1..J} C_j = \emptyset$ ,
     $\forall j \in \{1, \dots, J\}$  :
       $C_j = (\mathcal{V}_{C_j}, \mathcal{E}_{C_j})$  les points et relations associées à la cellule ;
  {ETAPE 2 : TRAITEMENT DES CELLULES DE VORONOI}
  for all  $j \in \{1, \dots, J\}$  do
    Réduction de dimension des profils de connectivité associés à  $C_j$ 
    Clustering de  $C_j$  d'après les profils réduits
    → Création de  $\cup_{k=1..K_j} A_k$  partition de  $\mathcal{V}_{C_j}$ 
    Sélection des parcelles valides ne touchant pas les frontières de la
    cellule dans

     $Parcelles_j = (\mathcal{V}^{c_j}, \mathcal{L}^{c_j})$  avec :  $\begin{cases} \mathcal{V}^{c_j} \text{ liste des points concernés,} \\ \mathcal{L}^{c_j} = l_1^j, \dots, l_{K_j}^j \text{ liste des labels associés.} \end{cases}$ 

    {AJOUT DANS LA PARCELLISATION FINALE DES PARCELLES VALIDES}
    Mise à jour de  $\mathcal{V}_{parcellisation}$  :
     $\mathcal{V}_{parcellisation} = \mathcal{V}_{parcellisation} \cup (\cup_{j=1..J} \mathcal{V}^{c_j})$ 
    Renumérotation des labels des  $Parcelles_j$  telle que
    Chaque parcelle aie un label unique
     $\min_k l_k^j = \mathcal{L}^{c_j}$ 
    Ajout de à  $\mathcal{L}_{parcellisation}$ 
    Mise à jour du domaine à parcelliser restant :
     $\mathcal{V}_{i+1} = \mathcal{V}_i \setminus \mathcal{V}_{parcellisation}$ 
  end for
   $i = i + 1$ 
end while

```

6.2.2 Traitement de chaque cellule

Une fois l'ensemble $\cup_{j=1..J} C_j$ obtenu, on traite chaque cellule C_j séparément selon les étapes suivantes :

1. Réduction de dimension des profils de connectivité,
2. Clustering des profils réduits,
3. Sélection des parcelles valides,
4. Mise à jour de la parcellisation finale.

6.2.2.1 1. Réduction de dimension des profils de connectivité

Les profils de connectivité de la cellule sont réduits en utilisant un ensemble de régions d'intérêt déterminées à partir du profil de connectivité total de la cellule et correspondant aux principales connexions rassemblées en faisceaux (cf. partie 5.2 et Fig. 5.3 pour plus de détails). Ces profils sont ensuite normalisés (norme L_2). Il faut noter que cette étape peut être remplacée par un autre type de réduction de dimension (analyse en composantes principales par exemple).

6.2.2.2 2. Clustering des profils réduits

La cellule est ensuite parcellisée à partir de ces profils de connectivité. L'algorithme utilisé pour le clustering est le **k-médoides** (Kaufmann et Rousseeuw 1990), appelé aussi **PAM** pour **Partition Around Medoids** dans le langage R^1 (c'est ce langage qui a été utilisé dans cette étape). Nous allons présenter ici seulement les principes de cette méthode (cf. article Kaufmann et Rousseeuw (1990) pour plus de détails sur l'algorithme).

La méthode des k-médoïdes Cet algorithme a pour but d'identifier k objets représentatifs appelés **médoïdes** parmi les objets d'un jeu de données.

Si l'on considère un ensemble de N observations $Obs = \{o_1, \dots, o_N\}$, et une mesure de dissimilarité d entre ces observations, l'objectif de cet algorithme est de trouver un sous-ensemble $\{m_1, \dots, m_k\}$ de Obs qui minimise la somme des dissimilarités entre les observations et l'observation représentative la plus proche. La fonction à minimiser est donc :

$$\sum_{i,N} \min_{t=1,\dots,k} d(i, m_t). \quad (6.1)$$

Cet algorithme peut être comparé à celui des **k-moyennes** ou k-means décrit précédemment (cf. Algorithme 1) où le centre de chaque cluster est défini par la moyenne des observations constituant la classe.

Le k-médoïdes a l'avantage d'être **plus robuste que le k-moyennes**. En effet, comme le centre d'un cluster est l'objet positionné de façon la plus centrale au sein du cluster et non la moyenne des objets (dans le cas du k-moyennes), le k-médoïdes est moins sensible aux données atypiques ainsi qu'à l'initialisation.

¹Logiciel libre spécialisé en statistique. <http://www.r-project.org>

Choix du nombre de clusters Le nombre de clusters K qui maximise la **largeur de silhouette moyenne** (*average silhouette width* en anglais) est choisi (Kaufmann et Rousseeuw 1990)). Ce critère permet de déterminer dans quelle mesure les observations (ici les profils de connectivité) au sein d'un cluster sont homogènes, et bien séparées des observations des autres clusters. Ainsi, si l'on considère un ensemble de N observations qu'on a regroupées en clusters, pour chaque observation j (correspondant à un point de la cellule), la **largeur de silhouette** $s(j)$ est définie comme suit :

Soit $a(j)$ la dissimilarité moyenne entre j et tous les autres points du cluster auquel j appartient (si j est la seule observation de ce cluster, $s(j) = 0$).

Pour tous les autres clusters C , soit $d(j, C)$ la dissimilarité moyenne de j avec toutes les observations de C .

La plus petite de ces $d(j, C)$ est $b(j) = \min_C d(j, C)$, et peut être interprétée comme la dissimilarité entre j et son cluster "voisin", i.e. le plus proche cluster auquel j n'appartient pas. Finalement,

$$s(j) := (b(j) - a(j)) / \max(a(j), b(j)).$$

Des observations avec un grand $s(j)$ (proche de 1) sont très bien clusterisées, tandis qu'un $s(j)$ faible (autour de 0) signifie que l'observation se situe entre deux clusters, et des observations avec un $s(j)$ négatif sont probablement placées dans un mauvais cluster.

6.2.2.3 Etapes 3. Sélection des parcelles valides et 4. Mise à jour de la parcellisation finale

Les parcelles sont considérées comme valides si elles ne sont pas à la frontière de la cellule. Les frontières d'une cellule sont constituées de l'ensemble des points de cette cellule qui sont juxtaposés à une autre cellule (Deux points sont considérés comme juxtaposés s'ils forment à eux deux une arête d'un polygone du maillage). Ainsi, les points de la cellule qui sont juxtaposés avec des points de régions déjà parcellisées (lors d'une autre itération) ne sont pas considérés comme des frontières ici.

Toutes les parcelles valides sont sauvegardées dans la parcellisation finale et enlevées de la surface restant à parcelliser qui devient S_{i+1} qui sera donc définie de la façon suivante : $\mathcal{V}_{i+1} = \mathcal{V}_i \setminus \cup \text{Parcelles Valides}$

6.2.3 Test d'arrêt

Si S_{i+1} est vide, l'algorithme s'arrête. Sinon, sous réserve que le nombre d'itérations ne dépasse pas un nombre maximal, on réitère le processus avec S_{i+1} . On se limite à 100 itérations au maximum, sachant qu'en pratique l'algorithme stoppe de lui-même au bout d'une dizaine d'itérations.

6.3 ESSAIS SUR LES DONNÉES SIMULÉES

6.3.1 Comparaison des parcellisations

Mesure : le Rand Index ajusté Nous allons utiliser le Rand Index ajusté pour comparer deux parcellisations (Adjusted Rand Index en anglais). Cet

index mesure combien de fois, pour une paire de points, les deux parcellisations A et B sont en accord (Hubert et Arabie 1985, Bohland et al. 2009a).

En notant :

- N_{11} le nombre de fois qu'un couple de points est dans le même cluster dans A ainsi que dans le même cluster dans B ,
- N_{00} le nombre de fois que pour un couple de points, chaque point est dans un cluster différent dans A ainsi que dans B ,
- N_{01} le nombre de fois où chaque point du couple est dans un cluster différent dans A , mais dans le même cluster dans B ,
- N_{10} idem que N_{01} en inversant A et B ,

Le Rand Index est la probabilité d'avoir les deux premiers cas :

$$RandIndex(A, B) = \frac{N_{11} + N_{00}}{N_{11} + N_{00} + N_{01} + N_{10}} \quad (6.2)$$

Le **Rand Index ajusté** est l'index précédent mais corrigé pour la chance.

Cette mesure a l'avantage de ne pas nécessiter d'appariement entre les parcelles des deux parcellisations concernées.

6.3.2 Cas optimal

Rappels Nous nous plaçons dans un premier temps dans le cas optimal (cf. partie Simulation 4.3 et Fig 5.7), avec :

- une densité de tracts élevée ($10 \text{ tracts}/\text{mm}^2$),
- la distribution locale des connexions d'une parcelle centrée autour du centroïde avec $\sigma = 5 \text{ mm}$ et $\gamma = 0.5$.

Avec ces données, la méthode fonctionne très bien et retrouve les parcellisations attendues. Dans un deuxième temps nous testerons le comportement de la méthode sur des cas moins optimaux (densités de tracts moins élevée, non gaussiannité de la répartition des connexions au sein de chaque parcelle).

Sujet modèle Dans la partie 5.2.3 nous avons vu que l'étape de réduction de dimension fonctionnait très bien avec ces données. Il n'est pas surprenant que ce soit aussi le cas pour le clustering. En effet, nous retrouvons une parcellisation à l'identique de celle qui a servi à générer les données. Le tableau de la figure 6.2 donne un récapitulatif des valeurs des silhouettes trouvées pour chaque parcelle ainsi que le Rand index ajusté.

Sur 5 sujets de la base simulée $B_{T_1}(10)$ Rappelons que les données simulées des sujets de cette base ont les mêmes caractéristiques que le sujet modèle T_1 précédent, avec une variabilité au niveau des parcelles de chaque sujet (les centroïdes de chaque parcelle étant décalés de 10 mm maximum par rapport au sujet modèle).

6.3.3 Cas extrême

Rappels Le sujet *modèle* du cas extrême des données simulées a pour caractéristiques :

- une connexion seulement par parcelle,

Sujets	iterations number	Cluster Silhouette Width			Average Silhouette	Adjusted Rand Index
		min	max	median		
modèle	4	0.72	0.93	0.89	0.89	1.0
sujet 01	4	0.82	0.95	0.89	0.90	1.0
sujet 02	4	0.56	0.94	0.86	0.87	1.0
sujet 03	4	0.56	0.97	0.91	0.88	1.0
sujet 04	4	0.61	0.97	0.91	0.90	1.0
sujet 05	4	0.75	0.96	0.92	0.90	1.0
A travers les sujets						
Moyenne	4	0.69	0.95	0.89	0.89	1.0
Ecart type	0	0.09	0.01	0.01	0.01	0.

FIG. 6.2 – Les silhouettes obtenues pour des données simulées “optimales” (sujet modèle T_1) ainsi que le Rand Index ajusté qui mesure le degré de similarité entre la parcellisation obtenue et celle attendue. Ici les deux sont identiques.

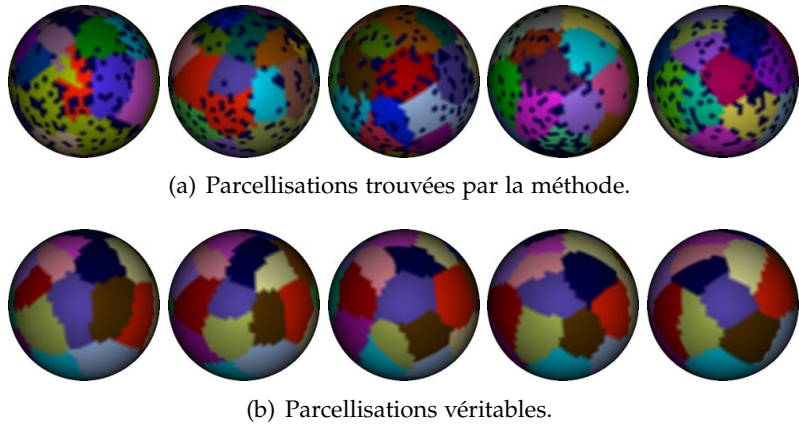


FIG. 6.3 – Cas extrême. Clustering du cortex entier des données simulées de 5 sujets.

- une densité de tracts faible (0.3 tracts/mm^2),
- la distribution locale des connexions d’une parcelle est moins gaussienne (autour du centroïde) avec $\sigma = 5 \text{ mm}$ et $\gamma = 0.9$.

A partir de ce *modèle*, une base de sujets avec une variabilité assez importante ($d_{max} = 20 \text{ mm}$) est construite : $B_{extreme}$.

Résultats Les parcellisations obtenues sont moins bonnes que dans le cas optimal précédent, avec en moyenne un Rand index ajusté de 0.51 avec comme écart type 0.08. Les résultats visualisés ne sont cependant pas catastrophiques comme le montre la figure 6.3.

6.4 ESSAIS SUR LES DONNÉES RÉELLES

Lors d’une autre étude (Roca et al. 2009a), j’ai parcellisé le cortex entier pour quatre sujets (dans un premier temps) en utilisant la méthode itérative décrite précédemment au 6.2.

Données et prétraitements L’étude a été effectuée sur un sous-ensemble de quatre sujets de la base NMR (cf. partie 2.1.2 pour une description précise). Les tracts ont été obtenus à partir du même tracking déterministe que précédemment 2.4.1. Les tracts sont ensuite filtrés par rapport à leur longueur : tous les tracts dont la longueur est inférieure à 10 *cm* sont éliminés. Ce seuil peut sembler élevé. Nous avons ici fait ce choix en faisant l’hypothèse que les tracts longue distance étaient plus reproductibles à travers les individus que les tracts courts (fibres en U par exemple). Les histogrammes des longueurs de tracts réalisés (cf. Fig. 6.4) nous montrent de plus qu’il reste une partie non négligeable des tracts.

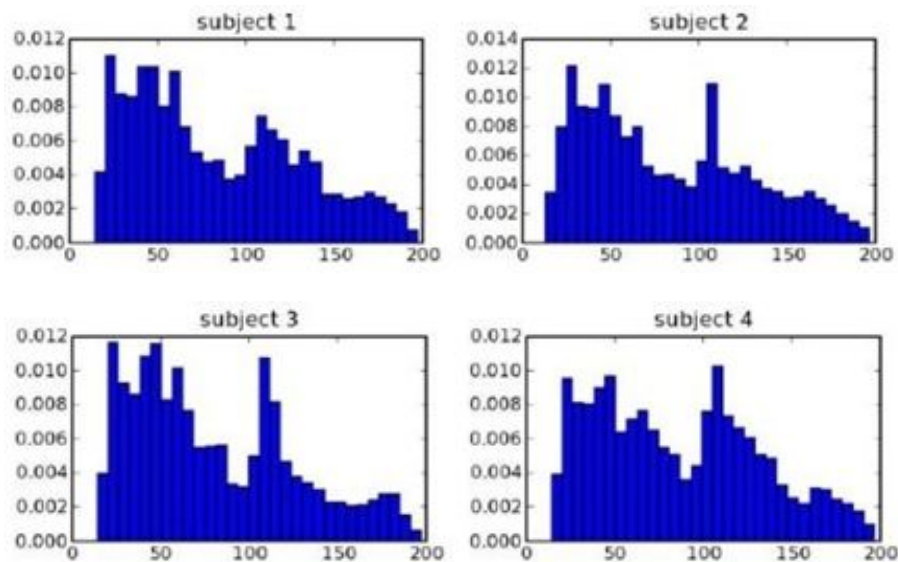


FIG. 6.4 – Histogramme des longueurs en mm des tracts pour les 4 sujets de l’étude Roca et al. (2009a).

6.4.1 Résultats

Exemple de parcellisations obtenues Lorsque l’on applique la méthode itérative pour parcelliser l’ensemble du cortex, on trouve en moyenne 140 parcelles par hémisphère. La Figure 6.5 montre la parcellisation obtenue pour un sujet, tandis que la figure 6.7(a) montre les parcelles obtenues autour de la région de Broca pour quatre sujets. Sur la figure 6.7(b) notamment, sont représentées les matrices de distance euclidienne entre les points de quelques parcelles. Les matrices ont été réordonnées par parcelle. L’organisation en “carré” de ces matrices est rassurante : pour chaque sujet indépendamment, le clustering trouve des parcelles relativement homogènes et séparées les unes des autres.

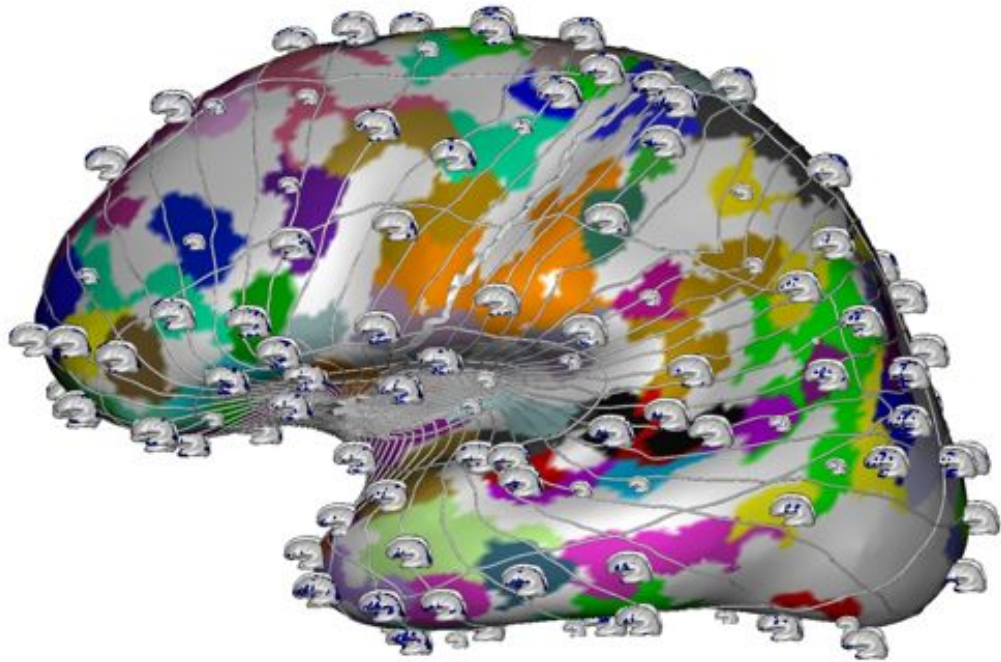


FIG. 6.5 – Résultats de la parcellisation intra-sujet : Les parcelles obtenues et leur profil de connectivité pour un des sujets. La visualisation est du type mixte régions-points (cf. 3.4). La grille représente un système de coordonnées sphériques construit à partir des sillons (Clouchoux et al. 2005).

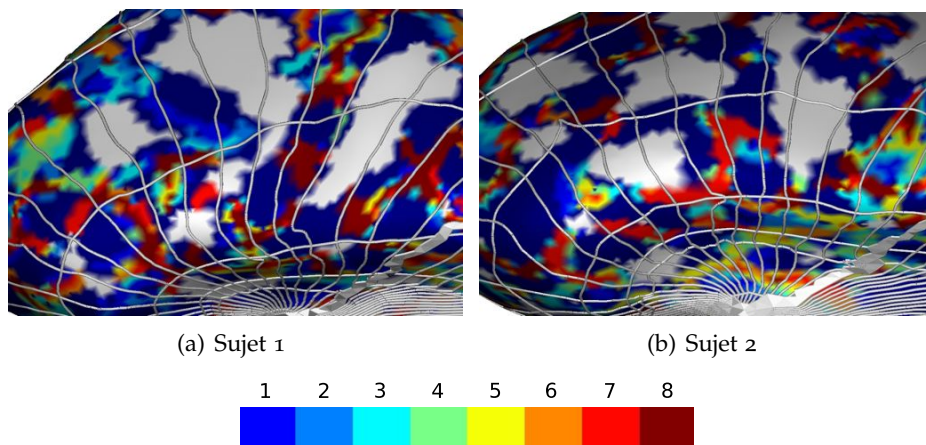
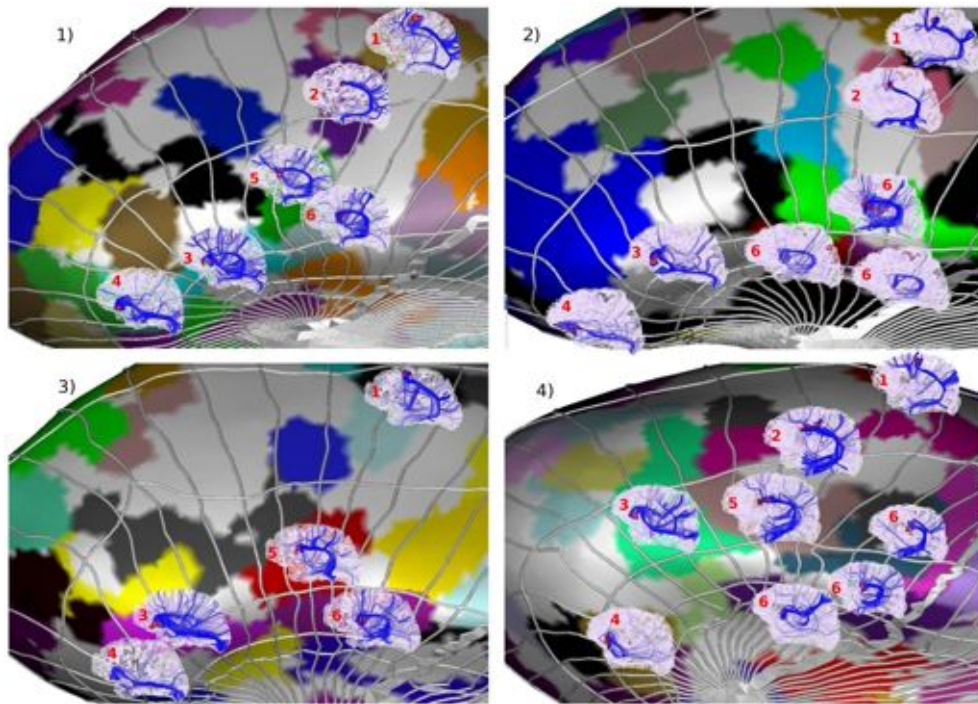


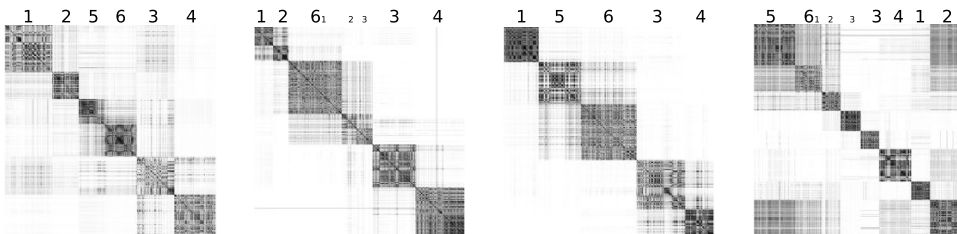
FIG. 6.6 – Sensibilité à l'initialisation. Cartes des frontières entre les parcelles trouvées sommées pour 8 parcellisations pour deux sujets, en faisant varier la division initiale du domaine. Zoom autour de la région de Broca. La palette indique le nombre de fois qu'un point est un point frontière.

Sensibilité à l'initialisation Pour évaluer la sensibilité du résultat final par rapport à l'initialisation (division du domaine et indirectement l'algorithme de clustering des k-médoïdes), nous avons effectué 8 fois la parcellisation en faisant varier la parcellisation de Voronoï centroïdale initiale. La figure 6.6 montre ainsi une carte des frontières obtenues entre les parcelles à travers ces différentes parcellisations pour le sujet 1, aux alentours de la région de Broca. Ce genre de carte permet ainsi d'avoir une idée des endroits du cortex où les frontières entre les parcelles sont stables à travers les différents essais.

Autour de Broca elles paraissent assez stables mais ce n'est pas le cas partout (cf. Fig. 6.6(a) juste au-dessus de la région de Broca.).



(a) Parcelles autour de Broca pour 4 sujets. Les parcelles sont numérotées selon la similitude de leur profil de connectivité à travers les sujets.



(b) Pour les 4 sujets : matrices de distance euclidienne entre les points des parcelles numérotées précédemment.

FIG. 6.7 – Résultats de la parcellisation intra-sujet pour 4 sujets : 6.7(a) Zoom sur la région de Broca pour 4 sujets. Une image représentant les tracts de chaque parcelle est rajoutée. 6.7(b) Matrices de distance euclidienne entre les points des parcelles numérotées précédemment. L'organisation en carrés reflète le fait que les parcelles sont homogènes et bien séparées les unes des autres en terme de profil de connectivité.

Cohérence des parcellisations entre les sujets Cependant, définir une correspondance entre les parcelles à travers les sujets est difficile. Sur la figure 6.7, j'ai attribué un numéro aux parcelles à travers les sujets selon la similitude de leur profil de connectivité. Il faut cependant noter que cette correspondance est mise ici à titre d'exemple et correspond juste à un essai d'appariement visuel par observation des tracts associés à chaque parcelle. Cette comparaison est rendue possible par l'outil de visualisation développé dans ce but (cf. 3.4.2).

6.5 CONCLUSION

Nous avons proposé dans ce chapitre une méthode pour parcelliser le cortex entier d'un sujet de manière itérative. Il est cependant difficile de trouver des correspondances évidentes entre les parcelles à travers les sujets. Pour résoudre ce problème, nous pouvons envisager au moins deux possibilités.

1. Une première option serait d'effectuer un **deuxième clustering inter-sujets** regroupant les parcelles individuelles au niveau du groupe de sujets. Pour cela, il faut calculer une distance entre deux parcelles de deux sujets différents. Un système de coordonnées surfaciques inter-sujets paraît approprié dans ce contexte. Un tel système mettant en correspondance les cortex des sujets fournit une base de décomposition commune et permet ensuite le calcul de différentes mesures de distance entre les profils de connectivité au niveau du groupe de sujets.
2. Une autre approche consiste à **réduire les profils** de connectivité et à effectuer le **clustering directement à l'échelle du groupe**. C'est cette approche que nous étudierons par la suite. Cela permet de parcelliser à partir d'informations de connectivité plus reproductibles à travers les sujets et la correspondance inter-sujets des parcelles est immédiate.

Après avoir ainsi élaboré et testé sur des données réelles la méthode itérative intra-sujet présentée dans ce présent chapitre, nous avons opté pour une analyse au niveau du groupe de sujets suivant la deuxième option, délaissant volontairement une validation plus poussée de la parcellisation intra-sujet. Cette analyse est présentée dans le chapitre suivant.

PARCELLISATION DE GROUPE DU CORTEX

7

CE chapitre est découpé en deux parties.

Une première partie présente une méthode parcellisant le cortex pour un groupe d'individus, aboutissant à une parcellisation spécifique à chaque individu avec une correspondance immédiate des parcelles à travers les individus. Pour ce faire, une réduction de dimension des données de connectivité au niveau du groupe est tout d'abord effectuée, de manière à sélectionner la connectivité pertinente (reproductible) au niveau du groupe. Un clustering est ensuite appliqué sur ces données réduites pour l'ensemble des sujets en même temps, soit en concaténant les données des sujets, soit en les moyennant. Ce travail a abouti à un article de conférence Roca et al. (2010a).

Dans une deuxième partie, nous comparerons sur des données simulées le comportement de cette méthode avec d'autres méthodes de parcellisation différant principalement par le type de réduction de dimension des données de connectivité utilisé, à savoir : soit le type précédent, soit à l'aide d'une analyse en composantes principales, soit à partir d'une segmentation a priori du cortex. Nous présenterons ensuite l'application de ces méthodes pour parcelliser la région englobant le SMA et le pré-SMA.

7.1 PARCELLISATION AVEC L'IDENTIFICATION DE RÉGIONS D'INTÉRÊT

Nous présentons ici une méthode parcellisant une région du cortex de plusieurs individus simultanément (un gyrus par exemple). Ce type d'approche inter-sujets a déjà été étudié dans la littérature. Guevara et al. (2008), dans une approche surfacique, propose ainsi une parcellisation inter-sujets du gyrus post-central de 3 sujets, en réduisant les profils de connectivité par rapport à une segmentation en gyri de la surface corticale, définie pour chaque sujet dans son espace natif.

Jbabdi et al. (2009), dans une approche volumique, réduit aussi les profils à partir d'une segmentation du cortex en 7 ou 48 régions, et en se plaçant dans un repère standard¹ (après recalage des sujets dans ce repère).

Nous allons présenter ici une démarche similaire à celles de Guevara et al. (2008) et Jbabdi et al. (2009), mais s'affranchissant d'une segmentation a priori du cortex pour réduire la dimension des données de connectivité. L'approche proposée s'appuie sur un système de coordonnées surfacique inter-sujets ainsi que sur la réduction de dimension des données de connectivité par identification de régions d'intérêt inter-sujets présentée dans la partie 5.2. Un clustering est ensuite effectué avec deux démarches possibles correspondant soit à une concaténation des données des sujets, soit à un moyennage.

Les hypothèses Cette démarche a pour but de parcelliser une région du cortex au niveau d'un groupe de sujets. Il faut donc d'ores et déjà supposer pouvoir identifier cette région :

- en intra-sujet : avec l'existence de critères pertinents et identifiables pour chaque sujet pour définir cette région corticale,
- en inter-sujets : les critères précédents doivent être relativement stables à travers les sujets. Nous verrons par la suite que le clustering selon l'approche par concaténation ne nécessite pas une correspondance des points de la région à travers les sujets. L'approche par moyennage fait en revanche cette hypothèse.

De plus, la réduction de dimension inter-sujets des données de connectivité s'appuie sur une relative reproductibilité des connexions à travers les sujets. Ainsi, nous supposons qu'il existe des tracts dont les terminaisons à l'approche du cortex sont relativement stables spatialement.

7.1.1 La méthode

Données d'entrée Pour chaque sujet, les données d'entrée sont les tracts représentant la connectivité anatomique issus d'un algorithme de tractographie, ainsi qu'une IRM anatomique. Nous rappelons que les prétraitements permettant d'obtenir ces données brutes sont détaillés au chapitre 2. Un résumé de ces traitements sera effectué dans la partie présentant les résultats sur les données réelles.

¹L'espace standard défini par le Montréal Neurological Institute (MNI).

Principe La méthode s’articule selon trois étapes présentées dans la figure 7.1 qui sont les suivantes :

1. une **Mise en correspondance inter-sujets des surfaces corticales** (avec Freesurfer) ;
2. une **Réduction de dimension inter-sujets** des données de connectivité sans a priori et sans recalage volumique, en utilisant la correspondance précédente et les données de connectivité elles-mêmes ;
3. un **Clustering inter-sujets** des points du gyrus à partir des profils réduits avec deux cas possibles :
 une **approche par concaténation**,
 et une **approche par moyennage**.

Nous allons décrire plus en détails chacune de ces étapes maintenant.

Etape 1. Analyse surfacique inter-sujet via freeSurfer Une analyse freeSurfer des images anatomiques des sujets (cf. Partie 2.3.2.2) nous fournit pour chaque sujet :

- la surface corticale $S = (\mathcal{V}, \mathcal{E})$ ayant une correspondance point à point avec celles des autres sujets ;
- une segmentation de cette surface en gyri grâce à une labélisation automatique de la surface corticale Desikan et al. (2006).

FreeSurfer fournit aussi une surface corticale moyenne utile pour la visualisation.

Les données utiles après les prétraitements sont les suivantes (cf.

Fig 7.1.1) :

- les surfaces corticales $(S^s)_{s=1..N}$ de l’ensemble des N sujets,
- les surfaces d’intérêt à parcelliser $(S^s_{gyrus})_{s=1..N}$, pour chaque sujet (on a choisi ici le gyrus post-central défini par la segmentation de freeSurfer), on a donc : $\forall s \ S^s_{gyrus} = (\mathcal{V}^s_{gyrus}, \mathcal{E}^s_{gyrus})$ où $\mathcal{V}^s_{gyrus} \subset \mathcal{V}^s$ et $\mathcal{E}^s_{gyrus} \subset \mathcal{E}^s$, avec $V^s_{gyrus} = \text{card}(\mathcal{V}^s_{gyrus})$
- les tracts de chaque sujet.

En pratique pour simplifier, on reportera la segmentation d’un sujet sur les autres sujets, pour avoir des points V^s_{gyrus} qui se correspondent à travers les sujets et du même nombre.

Etape 2. Construction de la matrice de connectivité La matrice de connectivité corticale de chaque sujet est construite suivant la méthode décrite dans la partie 3.3.2. On en déduit ensuite la matrice de connectivité du gyrus post-central $C \in \mathcal{M}_{(V_{gyrus}, V)}(\mathbb{R})$ pour chaque sujet. Cette matrice $C \in \mathcal{M}_{(V_{gyrus}, V)}(\mathbb{R})$ associée à S_{gyrus} définit par point $v_i \in \mathcal{V}_{gyrus}$ un profil de connectivité quantifiant la force de connexion de v_i vers chaque point $v_j \in \mathcal{V}$.

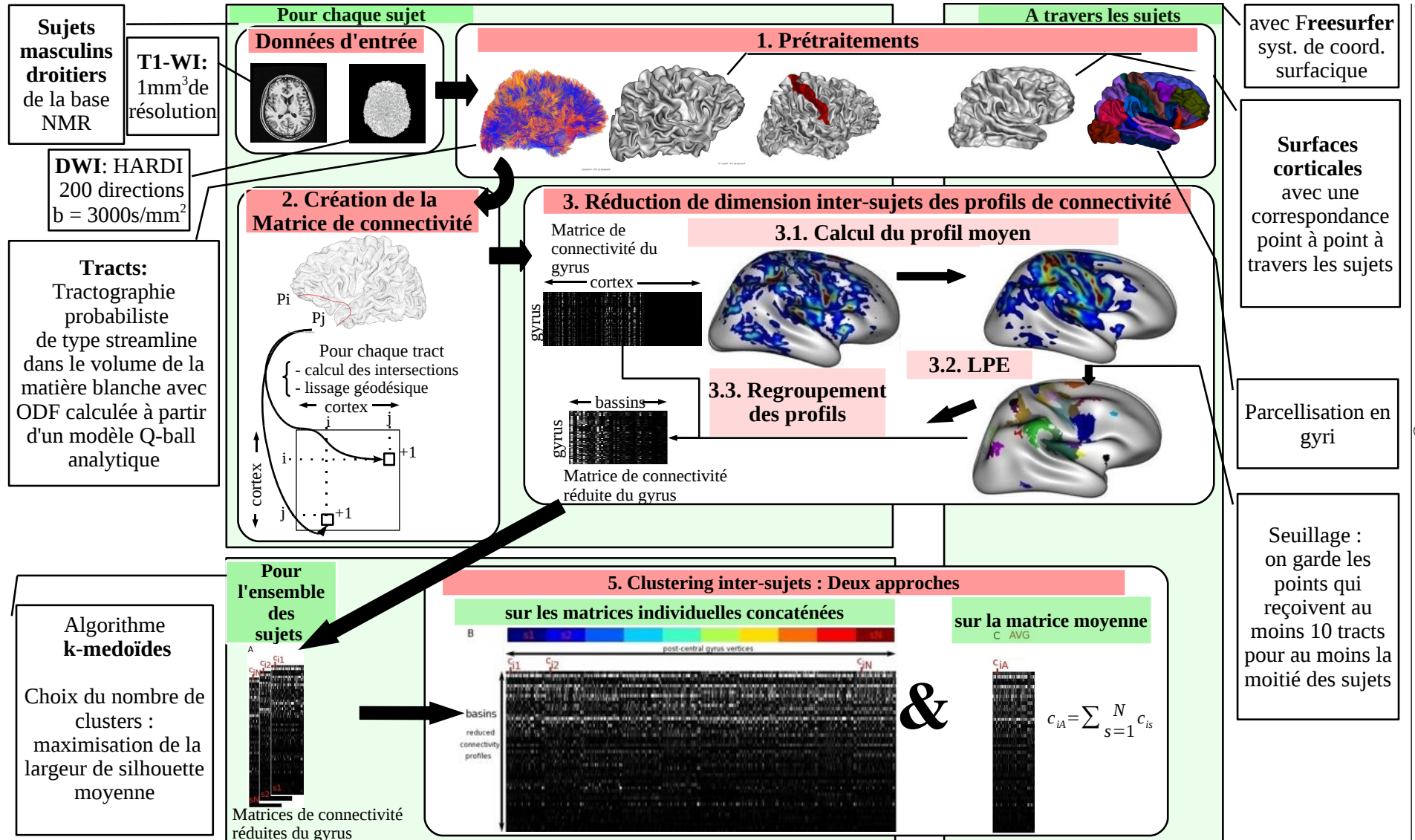


FIG. 7.1 – Réduction de dimension des profils de connectivité et clustering inter-sujets.

Etape 3. Réduction de dimension inter-sujets des données J'ai utilisé ici la réduction de dimension inter-sujets décrite au 5.2.2. Elle peut se résumer en quatre étapes :

1. Sujet par sujet : Calcul du profil de connectivité du gyrus ;
2. A travers les sujets : Moyenne des profils de connectivité du gyrus,
3. Sur le profil moyenné : Extraction des P régions (appelées ROIs) fortement connectées au gyrus par ligne de partage des eaux,
4. Pour chaque sujet s : Regroupement des profils de connectivité par rapport à ces régions : pour chacune des P régions précédentes, on somme les connexions des points correspondants. On a alors une matrice de connectivité réduite de taille (V_{gyrus}^s, P) .

Etape 4. Clustering des points du gyrus à partir des profils Les profils de connectivité de tous les sujets sont ensuite normalisés (norme L_2) et un clustering des points du gyrus est effectué à partir de ces profils, pour l'ensemble des sujets, selon deux stratégies.

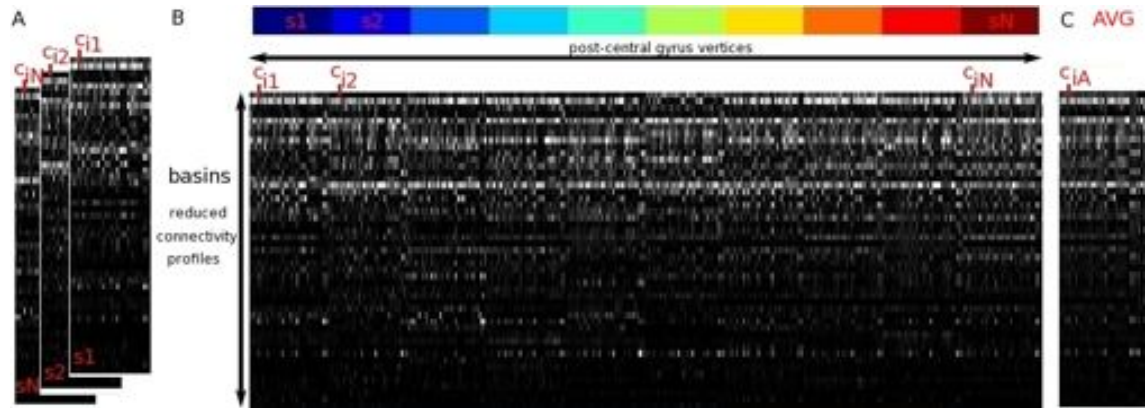


FIG. 7.2 – Profils de connectivité des points du gyrus post-central regroupés : A. par sujet ; B. concaténés pour l'ensemble des sujets ; C. Moyennés à travers les sujets ($c_{iA} = \frac{1}{N} \sum_{s=1}^N c_{is}$).

Concaténation des sujets Dans un premier temps, on concatène les profils de connectivité des points du gyrus post-central de l'ensemble des sujets (cf. Fig. 7.2.B) dans une grande matrice (de taille $(N * V_{gyrus}, P)$). Un clustering est ensuite appliqué à cette matrice sans se soucier des sujets sous-jacents. Le clustering utilisé ici est le kmedoids (appelé PAM dans le langage R Kaufmann et Rousseeuw (1990)). La mesure de dissimilarité associée est la distance euclidienne entre les profils. Le nombre de clusters supérieur à deux et maximisant la largeur de silhouette moyenne (cf. définition à la partie 6.2.2.2) est choisi. On obtient ainsi une parcellisation par sujet avec une correspondance directe entre les parcelles des différents sujets car le clustering a été fait simultanément sur tous les sujets. Cette méthode a cependant des inconvénients, notamment au niveau de la dimension. En effet en concaténant les sujets, on obtient des matrices de grande taille, et le calcul en est ralenti.

Moyenne des sujet Une autre approche consiste à moyenner les profils de connectivité à travers les sujets (cf. Fig. 7.2.C), pour ensuite clusteriser les points du gyrus du “sujet moyen” ainsi créé. Nous pouvons faire cela car nous avons une correspondance point à point à travers les sujets du domaine à parcelliser. En effet, à chaque point i_A du gyrus (sur le cortex moyen), N points $\{i_1, \dots, i_N\}$ correspondant aux N sujets sont associés, et on peut ainsi calculer le profil moyen c_{iA} en chaque point avec : $c_{iA} = \frac{1}{N} \sum_{s=1}^N c_{is}$.

On peut ensuite effectuer le clustering à partir de ces profils moyennés. Cette deuxième approche est plus rapide car la dimension du problème est réduite en moyennant.

Il est intéressant de combiner ces deux approches car nous allons pouvoir comparer leurs résultats. Dans la première, nous pourrions identifier des parcelles spécifiques à quelques sujets, alors que dans la deuxième, seulement les parcelles reproductibles à travers les sujets vont apparaître.

Dans la suite, nous allons présenter le comportement de cette méthode sur des données simulées, puis dans le cas de données réelles.

Notations Nous nous servirons des notations suivantes :

Soit $\mathcal{S} = s_1, \dots, s_S$ un groupe de sujets. Pour un sujet s_i :

- $P_{True}^{s_i}$: Parcellisation gold standard du sujet s_i (cas des données simulées)
- $P_A^{s_i}$: Parcellisation du sujet s_i obtenue par la méthode A

Nous noterons P_{True}^S la parcellisation qui rassemble l'ensemble des sujets (par concaténation des $P_{True}^{s_i}$).

Pour évaluer la concordance entre les parcellisations obtenues et les parcellisations “gold standard”, le **Rand Index ajusté** sera utilisé. C'est une mesure de comparaison entre deux parcellisations ne nécessitant pas un appariement préalable entre les parcelles (cf. partie 6.3.1 pour une définition plus précise). Nous pouvons calculer cet index pour deux cas différents :

- soit par sujet s_i : $adjRandIndex(P_{True}^{s_i}, P_{wat}^{s_i})$: le calcul est fait indépendamment pour chaque sujet. Ainsi, si l'appariement des parcelles des $P_{wat}^{s_i}$ à travers les sujets n'est pas cohérent par rapport aux $P_{True}^{s_i}$, cette mesure n'y est pas sensible ;
- soit sur la concaténation des sujets : $adjRandIndex(P_{True}^S, P_{wat}^S)$: contrairement à la précédente, cette mesure tient compte de la cohérence des parcelles entre les sujets.

7.1.2 Essais sur les données simulées

7.1.2.1 Cas optimal

Sujets de la base simulée $B_{optimale}$ Rappelons que les données simulées des sujets de cette base ont les mêmes caractéristiques que le sujet modèle T_1 (densité élevée de tracts, distribution locale des connexions d'une parcelle centrée autour du centroïde), avec une variabilité au niveau des parcelles de chaque sujet (les centroïdes de chaque parcelle étant décalés

de 10 mm maximum par rapport au sujet modèle). Pour plus de détails, veuillez vous reporter à la description des données simulées 4.3. La parcellisation est effectuée sur le gyrus 1 de la segmentation en gyri du sujet modèle T_1 (cf. Fig. 5.7(a)). L'analyse est faite sur 10 sujets. Les étapes 1 à 3 ont déjà été décrites précédemment. Nous présentons ici les résultats de la quatrième étape : le clustering inter-sujets des profils réduits des points du gyrus, après leur concaténation pour l'ensemble des sujets, ou leur moyennage à travers les sujets.

Résultats Le nombre de clusters optimal trouvé est 10 clusters. Il permet de retrouver pour l'ensemble des sujets les 4 principales parcelles au centre du gyrus, ainsi que les parcelles à la frontière qui sont à l'extérieur du gyrus du sujet modèle T_1 mais comprises dans le gyrus des sujets de la base simulée dérivée (cf. Fig. 7.4(b)). Certaines de ces parcelles frontières ne sont cependant pas retrouvées (la parcelle bleue claire en bas à droite par exemple). Cela peut s'expliquer par le fait qu'elles n'ont pas un nombre suffisant de points.

Dans ce cas optimal, les indices de concordance sont élevés avec un Rand index ajusté > 0.98 .

	Adjusted Rand Index
$P_{True}^{S_i} / P_{wat}^{S_i}$	$0.99 \pm 0.3 \cdot 10^{-3}$
P_{True}^S / P_{wat}^S	0.98

FIG. 7.3 – Parcellisations de la base de sujets simulés $B_{optimale}$. Concordance avec les parcellisations Gold Standard de chaque sujet $P_{True}^{S_i}$. L'Adjusted Rand Index est calculé soit sujet par sujet, soit sur la concaténation des sujets.

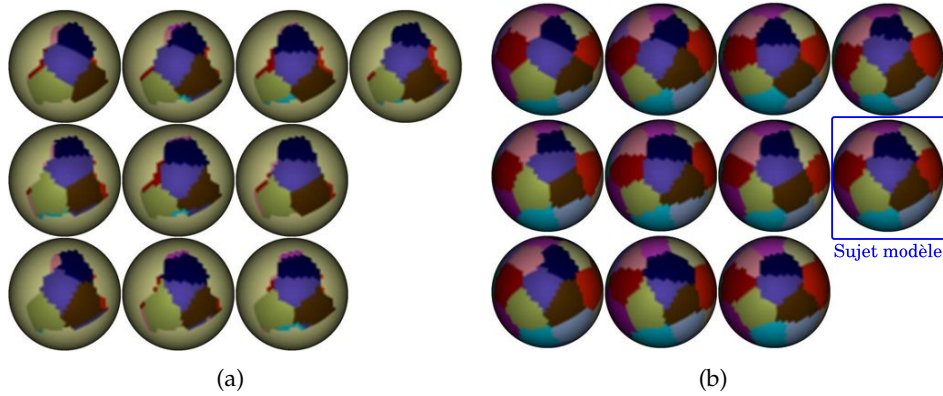


FIG. 7.4 – Base de sujets simulés $B_{optimale}$. 7.4(a). Parcellisations $P_{wat}^{S_i}$ obtenues en concaténant les sujets ; 7.4(b) Parcellisations simulées $P_{True}^{S_i}$ et parcellisation du sujet modèle T_1 (entourée en bleu).

7.1.2.2 Cas extrême

Rappels Le sujet modèle du cas extrême des données simulées a pour caractéristiques :

- une connexion seulement par parcelle,
- une densité de tracts faible (0.3 tracts/mm^2),
- la distribution locale des connexions d'une parcelle est moins gaussienne (autour du centroïde) avec $\sigma = 5 \text{ mm}$ et $\gamma = 0.9$.

A partir de ce modèle, une base de sujets B_{extreme} avec une variabilité assez importante ($d_{\text{max}} = 20 \text{ mm}$) est construite.

Dans cette base, le gyrus 1 a 4 parcelles principales, et d'autres parcelles secondaires qui dépendent de la variabilité de chacun des sujets par rapport au sujet modèle.

La figure 7.5 montre le résultat d'un clustering inter-sujets dans deux cas de figure :

- après réduction des profils de connectivité au niveau du groupe de sujets (cf. partie 5.2.4 pour visualiser les régions d'intérêt identifiées pour cette réduction de dimension), nous appellerons ce cas M_{ROIs} ,
- clustering à partir de la matrice de corrélation des profils préalablement réduits en fonction d'une segmentation de Voronoï centroïdale de la surface corticale simulée, nous appellerons ce cas $M_{\text{cross-corr}}$.

Résultats Le nombre optimal de clusters est de 2 dans le cas du clustering effectué sur la matrice de cross-corrélation et ne correspond pas dans ce cas au nombre de clusters attendus. Il est de 11 dans le cas de la réduction par identification de régions d'intérêt, avec 4 clusters principaux (cf. Fig. 7.5(b)) qui sont cohérents spatialement avec les clusters véritables des sujets (cf. Fig. 7.5(e)).

Pour un nombre de clusters fixé à 5, les clusters trouvés par M_{ROIs} sont plus satisfaisants que ceux identifiés par $M_{\text{cross-corr}}$. Dans la méthode $M_{\text{cross-corr}}$ il y a en effet une classe "poubelle" : du fait de la faible densité de connexions, les points d'une parcelle du gyrus ont des profils de connectivité complets différents au sein même d'un sujet et le phénomène est accentué à travers les sujets. Dans le cas de la méthode M_{ROIs} , les similarités entre les points d'une même parcelle sont plus élevées même dans ce cas, comme nous l'avons vu dans le chapitre 5 sur la réduction des profils de connectivité, en visualisant les matrices de corrélation entre les profils complets et non réduits (cf. Fig. 5.38). La figure 5.18 du même chapitre illustre également bien ce phénomène avec un exemple dans le cas des données réelles.

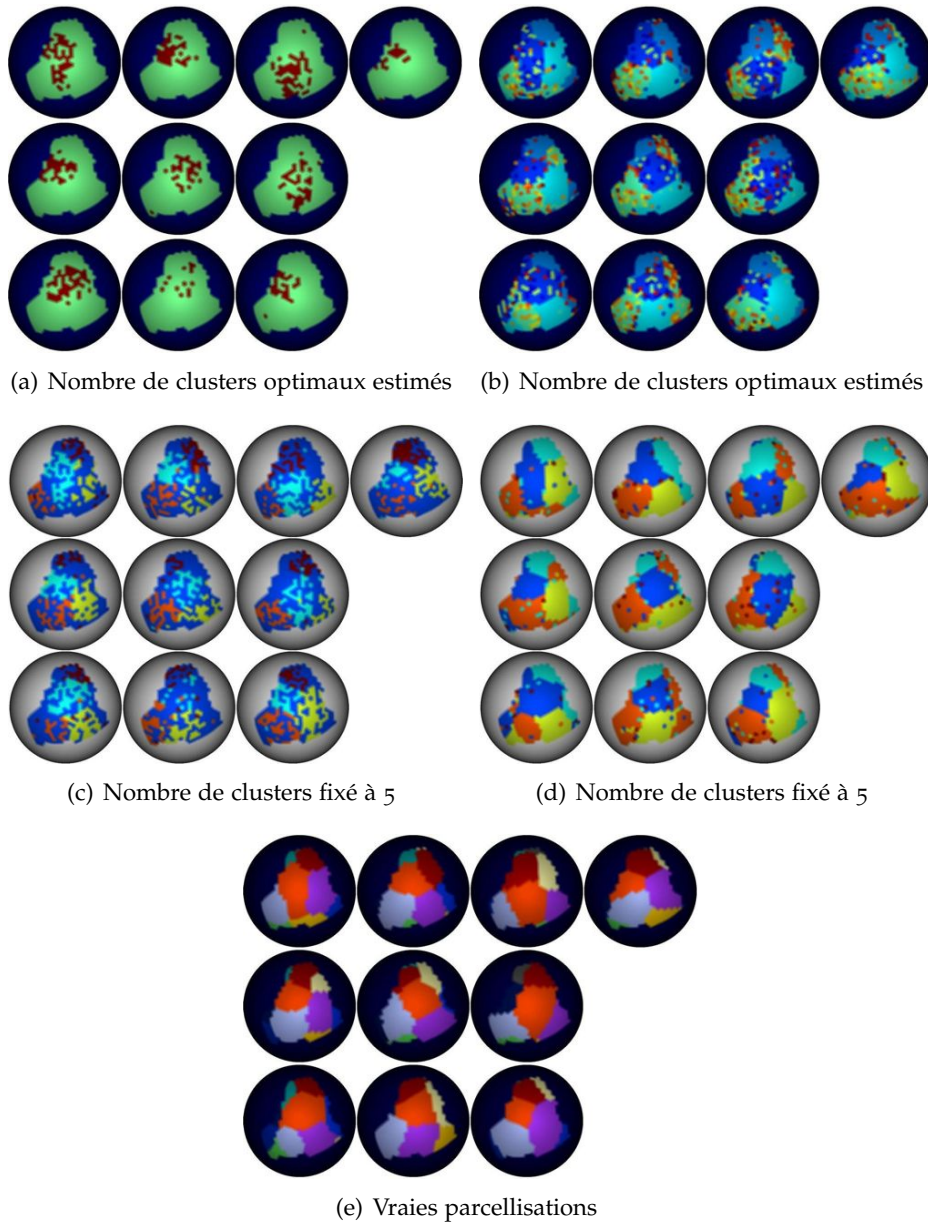


FIG. 7.5 – Clustering inter-sujets du cas extrême $B_{extreme}$ pour 2 cas : 7.5(a) et 7.5(c) Sur la matrice de cross-correlation (construite à partir d'un profil réduit par rapport à une parcellisation aléatoire en 200 parcelles), 7.5(b) et 7.5(d) Sur la matrice de connectivité réduite en inter-sujets décrite précédemment, 7.5(e) Parcellisations Gold standard des 10 sujets.

7.1.3 Essais sur les données réelles

Parcellisation du gyrus post-central

Données et prétraitements Les résultats présentés ici sont issus de l'article de conférence Roca et al. (2010a). Lors de cette étude, 10 sujets de la base NMR (présentée dans la partie 2.1.2) ont été analysés. Nous avons reconstruit les tracts grâce à une tractographie probabiliste de type "streamline" décrite en détail précédemment dans la partie 2.4.2 à partir des données HARDI de la base NMR. Les "Orientation Distribution

Functions" (ODF) ont été modélisées par un modèle analytique de type Q-ball (Descoteaux et al. 2007). L'analyse freesurfer des sujets a été réalisée par Alan Tucholka alors qu'il était en thèse au LNAO, aboutissant à des surfaces corticales de 80 000 points environ. La matrice de connectivité est calculée suivant la méthode décrite précédemment 3.3.2 et lissée géodésiquement par rapport à la surface (avec un noyau de lissage de 5 mm). Aucun filtrage par rapport à la longueur des tracts n'est effectué. L'idée étant que la dimension de réduction des profils au niveau du groupe permet de sélectionner les tracts reproductibles et donc que seulement les tracts courte distance pertinents seront gardés.

Parcellisations obtenues Les résultats obtenus sont assez encourageants. Les parcellisations obtenues à partir des deux approches (profils concaténés ou moyennés) sont en effet cohérentes, tant en termes de localisation des parcelles sur le cortex qu'en termes de profil de connectivité.

Pour l'approche moyennée, le nombre optimal de clusters trouvé est 5 (Fig. 7.8) alors qu'il est de 7 (Fig. 7.7) pour l'approche avec concaténation. Le code couleur choisi établit une correspondance entre les parcelles trouvées dans les deux approches. On retrouve des profils de connectivité comparables entre les parcelles de même couleur à travers les deux approches. Il y a deux parcelles additionnelles dans l'approche par concaténation des profils, une parcelle bordeaux et une parcelle rouge (cf. Fig. 7.6 et 7.7). La parcelle bordeaux se retrouve ainsi chez deux sujets. Nous n'avons pas vraiment d'explication pour cette parcelle, qui peut résulter d'un artefact dans les données de diffusion associées. La parcelle rouge a l'air elle de résulter du découpage de la parcelle vert pomme (du sujet moyenné) en deux parcelles. Comme elle varie en localisation, il paraît cohérent qu'elle disparaisse dans l'étude sur le sujet moyenné.

Dans le cas de l'approche avec concaténation, les parcelles ont un profil de connectivité propre qui est similaire entre les sujets et bien séparé des autres parcelles. De plus, on peut remarquer que dans l'approche concaténée, certaines parcelles existent seulement pour un sous-ensemble de sujets. Deux parcelles additionnelles sont trouvées par rapport au clustering sur la matrice moyennée. On peut alors envisager qu'avec un groupe de sujets hétérogène, on puisse mettre en évidence des parcelles spécifiques à un groupe (avec un groupe de patients et de témoins par exemple).

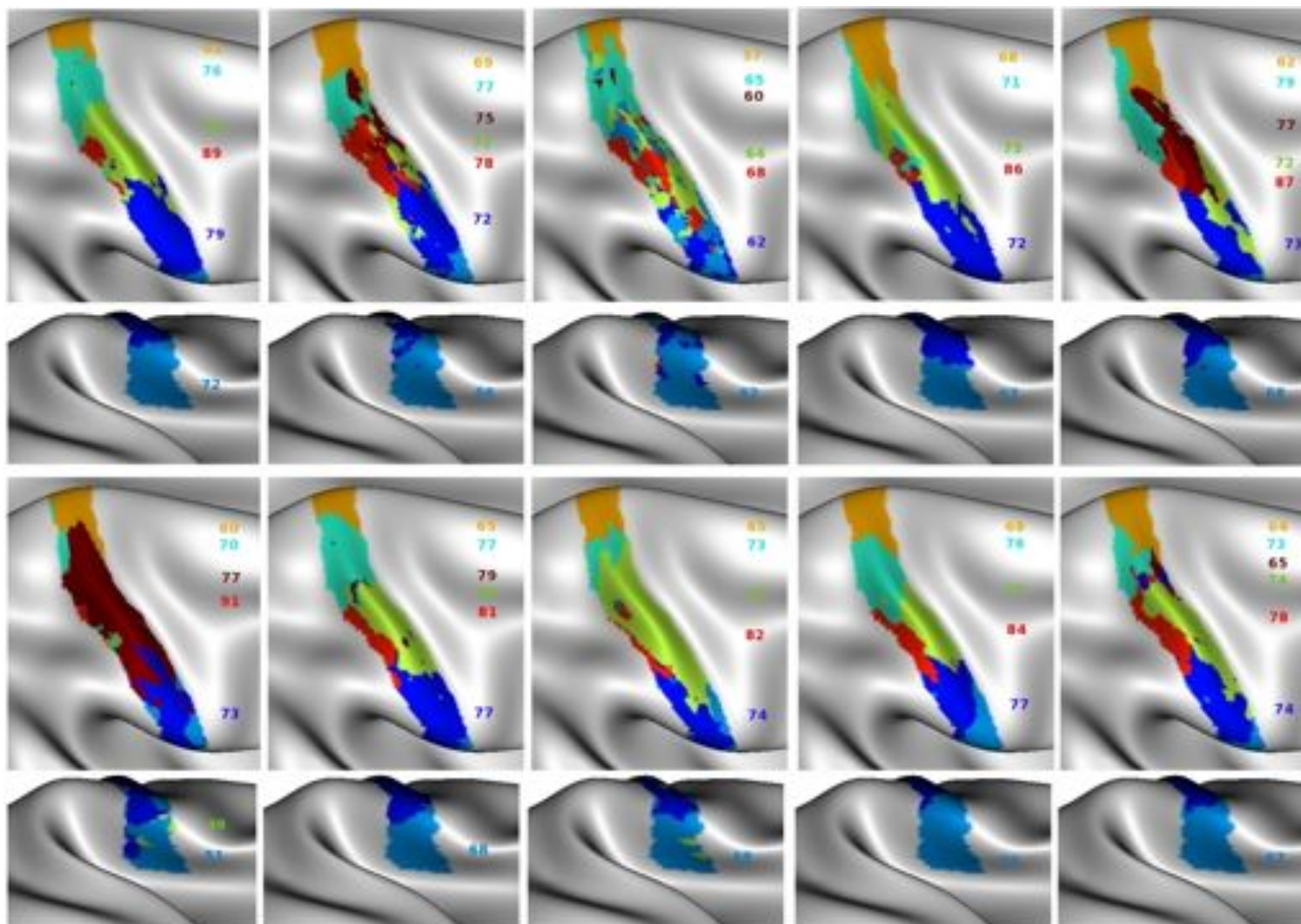


FIG. 7.6 – Clustering du gyrus post-central avec l'approche concaténée : Parcellisations obtenues pour chacun des 10 sujets projetées sur la surface corticale moyenne. Les chiffres représentent pour chaque sujet le pourcentage de tracts dans la connectivité de chaque parcelle ayant été prise en compte par le clustering.

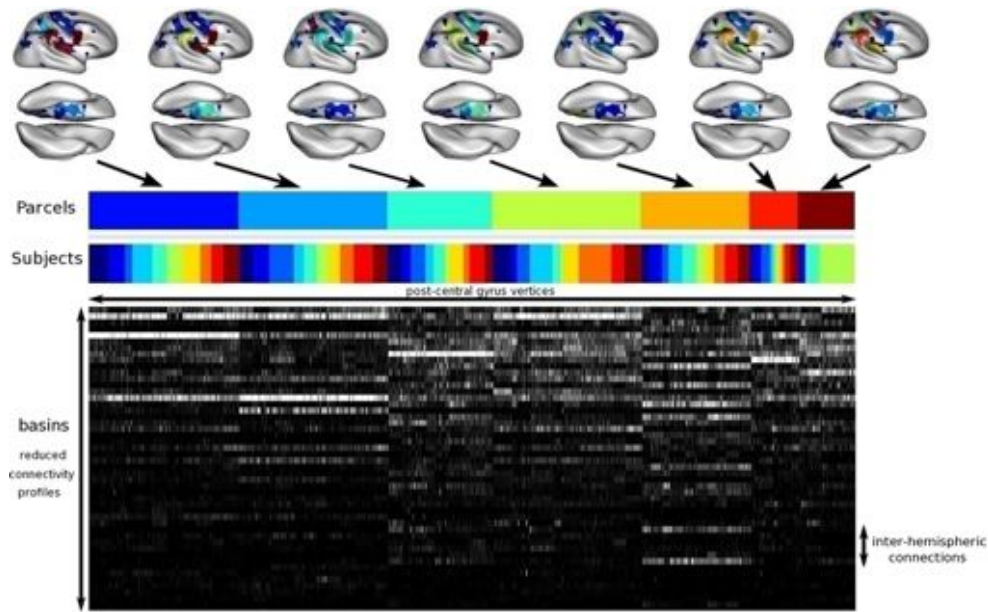


FIG. 7.7 – Clustering du gyrus post-central avec l'approche concaténée : Les profils de connectivité réduits de chacun des points du gyrus pour chacun des sujets (ceux de la figure 7.2.B) réordonnés en fonction des parcelles et des sujets.

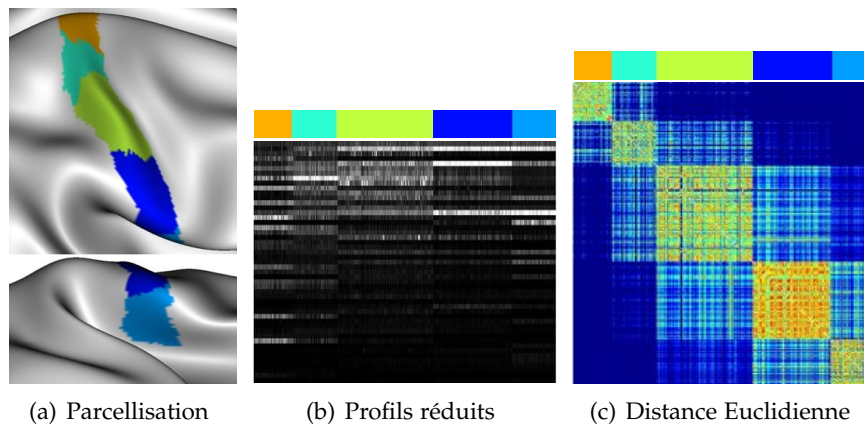


FIG. 7.8 – Clustering du gyrus post-central avec l'approche moyennée : 7.8(a). Parcélisation obtenue projetée sur la surface corticale moyenne ; 7.8(b). Les profils de connectivité moyens réordonnés en fonction des parcelles (les mêmes que ceux de la figure 7.2.C) ; 7.8(c). Matrice de distance euclidienne associée aux profils réduits.

7.1.4 Discussion

Nous avons démontré sur l'exemple du gyrus post-central, la **possibilité de parcelliser une région corticale pour un groupe de sujets sans utiliser de segmentation a priori** du cortex pour réduire les profils de connectivité avec l'obtention de **parcelles homogènes et relativement bien séparées** en terme de profils de connectivité réduits, ainsi que **cohérentes spatialement à travers les sujets**.

Sur les données simulées, nous obtenons de meilleurs résultats avec la réduction de dimension des profils plutôt qu'en effectuant un clustering directement à partir de la matrice de cross-corrélation entre les profils

complets (en fait réduits à l'aide d'une segmentation de Voronoï centroïdale en 200 cellules de la surface corticale simulée).

Dans la deuxième partie de ce chapitre, nous allons comparer cette méthode à d'autres méthodes (avec d'autres types de réduction de dimension des profils de connectivité) sur les données simulées et les données réelles.

7.2 COMPARAISON DE DIFFÉRENTES MÉTHODES

7.2.1 Récapitulatif des approches

Nous avons testé dans cette thèse trois autres méthodes pour résumer les profils de connectivité :

- la réduction à partir d'une **parcellisation a priori en gyri** (de FreeSurfer),
- l'**analyse en composantes principales**,
- le **calcul de la cross-correlation** entre les profils.

A chacun de ces types de réduction des profils correspond deux façons de parcelliser le cortex, en intra et en inter-sujets. La table 7.9 donne les notations utilisées pour nommer ces méthodes.

Notation	Méthode de réduction de dimension associée
M_{ROIs}	Identification automatique de régions d'intérêt
M_{PCA}	Analyse en composantes principales
M_{Gyri}	Segmentation a priori (parcellisation en gyri de FreeSurfer)
M_{Cross}	Calcul de la cross-correlation entre les profils

FIG. 7.9 – Les méthodes de parcellisation corticale utilisées en fonction du type de réduction de dimension des profils effectué. En fonction du caractère intra ou inter-sujets, un exposant sera rajouté (M_*^{intra} ou M_*^{inter}) s'il y a ambiguïté.

7.2.2 Procédure

Dans la suite, nous allons présenter le comportement de ces méthodes sur des données simulées, puis dans le cas de données réelles, pour la parcellisation corticale d'un patch du cortex (composé de P points et noté $\mathcal{V}_{patch} = \{v_1, \dots, v_P\}$).

Notations Nous nous servons des notations suivantes :

Soit $\mathcal{S} = s_1, \dots, s_S$ un groupe de sujets. Pour un sujet s_i :

- $P_{True}^{s_i}$: Parcellisation gold standard du sujet s_i (cas des données simulées) sous forme d'un vecteur de labels, un label étant associé à chaque point du patch, $P_{True}^{s_i} \in (\mathbb{N}^*)^P$
- $P_A^{s_i}$: Parcellisation du sujet s_i obtenue par la méthode A

Nous noterons P_{True}^S la parcellisation qui rassemble l'ensemble des sujets (par concaténation des $P_{True}^{s_i}$ à travers les sujets).

Mesure : Rand Index ajusté Nous allons utiliser le Rand Index ajusté pour comparer deux parcellisations (cf. partie 6.3.1 pour une définition).

Cet index mesure combien de fois, pour une paire de points, les deux parcellisations A et B sont en accord (Hubert et Arabie 1985, Bohland et al. 2009a). Il a l'avantage de ne pas nécessiter d'appariement entre les parcelles des deux parcellisations concernées.

Cohérence sujet / gold standard du sujet

Sujet par sujet Pour comparer, pour un sujet s_i , la parcellisation $P_A^{s_i}$ obtenue par la méthode A avec la parcellisation gold standard du sujet $P_{True}^{s_i}$, nous calculerons le Rand index ajusté entre ces deux parcellisations.

Pour l'ensemble des sujets Lorsque le clustering est effectué sur les profils concaténés à travers les sujets, nous obtenons une parcellisation pour la concaténation des points du patch à travers les sujets, pour une méthode A donnée. Cette parcellisation sera notée P_A^S et représente la concaténation pour l'ensemble des N sujets des $\{P_A^{s_i}\}_{i=1}^N$. Nous mesurerons le Rand index ajusté entre cette parcellisation P_A^S et la concaténation des parcellisations gold standard de chacun des sujets P_{True}^S . Cela nous permet ainsi d'évaluer la cohérence entre les parcelles des différents sujets.

7.2.3 Résultats sur données simulées

Description des données

Les données simulées ici correspondent à différentes bases de sujets $B_{T_{density}, \gamma}(d_{max})$ avec des paramètres variables :

- $d_{max} \in [15, 20, 30, 40]$
- $density \in \{0.2, 0.3, 0.4, 0.6, 0.8, 1\}$,

et fixes comme $\sigma = 5$ et $\gamma = 0.9$. Lors de la connexion du graphe binaire des connexions entre les parcelles, une seule connexion par parcelle est choisie (le degré de chacun des noeuds du graphe est 1).

Comparaison des méthodes

Peu de variabilité : $d_{max} = 15$ La figure 7.10 montre l'évolution des Rand index ajustés pour les 3 méthodes M_{ROIs} , M_{Gyri} , M_{PCA} en fonction de la densité des connexions. Pour chaque valeur de densité et chaque méthode, une moyenne des mesures est effectuée à travers les 6 gyri du cortex (courbes).

Cette figure permet de comparer les méthodes en fonction de la densité des connexions, ainsi que d'avoir un aperçu de la variabilité de ces méthodes à travers les gyri de la surface corticale simulée.

Différences en fonction de la densité des connexions Pour une densité < 0.4 , les méthodes M_{ROIs} et M_{Gyri} sont plus performantes que M_{PCA} :

- Quand il y a peu de tracts (densité inférieure ou égale à 0.4 tracts par mm^2 , la méthode M_{ROIs} est plus performante, vient ensuite celles des gyri M_{Gyri} , puis la PCA ;

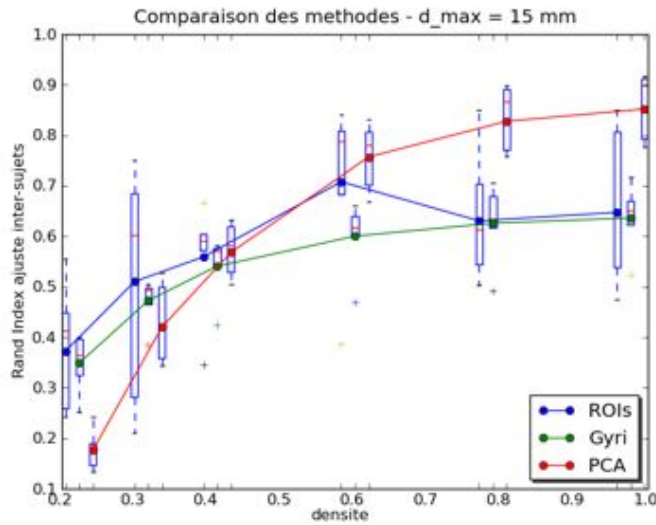


FIG. 7.10 – Sur 10 sujets. Comparaison des méthodes inter-sujets sur la base simulée $B_{T_{density, \gamma=0.9}}(d_{max} = 15mm)$ pour différentes valeurs de densité de connexion. Les courbes représentent la moyenne des rand index ajustés pour chaque méthode, calculée à travers les 6 gyri de la base.

- M_{PCA} dépasse M_{Gyri} et M_{ROIs} à partir d'une densité de connexions égale à 0.4 et 0.6 respectivement.
- la méthode M_{Gyri} est toujours inférieure en moyenne à M_{ROIs} .

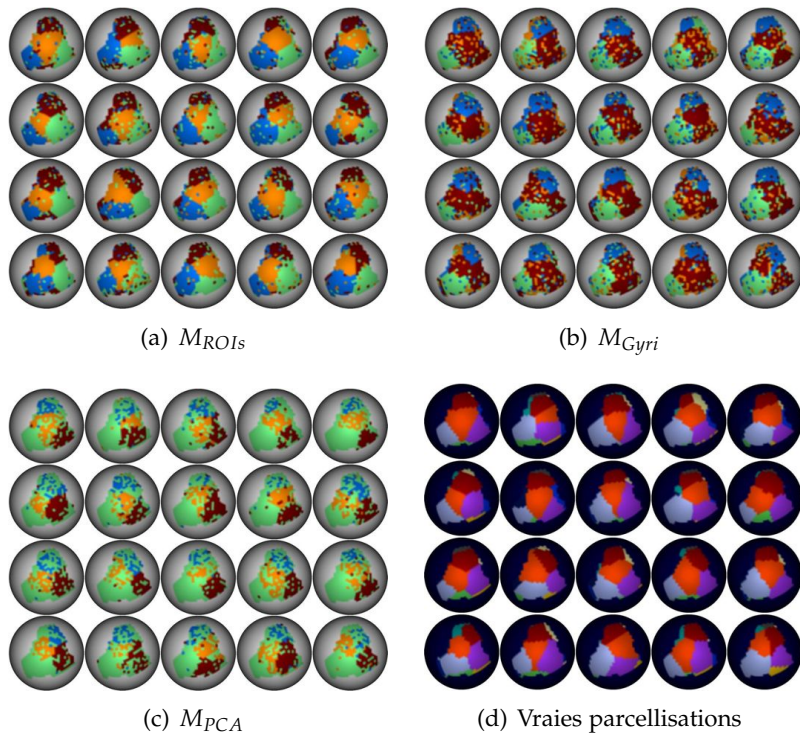


FIG. 7.11 – Exemples de parcellisations du gyrus 1 avec une variabilité inter-sujets $d_{max} = 15mm$ et une densité de connexion à 0.2. Clustering inter-sujets sur 20 sujets. 7.11(a), 7.11(b), 7.11(c). Parcellisations obtenues pour les trois méthodes M_{ROIs} , M_{Gyri} , et M_{PCA} avec un nombre de clusters fixé égal à 4. 7.11(d) Parcellisations gold standard du gyrus 1.

La figure 7.11 montre l'exemple de la parcellisation du gyrus 1 dans le cas d'une densité de connexion faible à 0.2 tracts par mm^2 .

Nous pouvons remarquer que :

- Avec la méthode M_{Gyri} : il n'y a pas de différenciation entre les parcelles orange et violette (des parcellisations gold standard des sujets). Ceci s'explique par le fait que les connexions de ces deux parcelles sont regroupées dans le même gyrus.
- Les points avec peu de connexions ne sont pas assez similaires pour être regroupés par la méthode M_{PCA} ,
- Avec la méthode des M_{ROIs} l'identification des régions d'intérêt se passe mieux car la somme des connexions est moyennée à travers les points du gyrus et à travers les sujets.

Variabilité à travers les gyri

- Il y a une variabilité importante entre les mesures obtenues à travers les gyri pour la méthode M_{ROIs} ;
- Les autres méthodes présentent beaucoup moins de variabilité (et M_{Gyri} en a très peu comparée aux deux autres).

Influence du nombre de sujets Lorsque l'on double le nombre de sujets sur lequel est effectué la réduction de dimension et le clustering des profils de connectivité, nous observons que :

- Le comportement des méthodes M_{Gyri} et M_{PCA} ne change pas,
- La variabilité de M_{ROIs} s'amointrit (cf. Fig. 7.12) et 7.13, et 7.14).

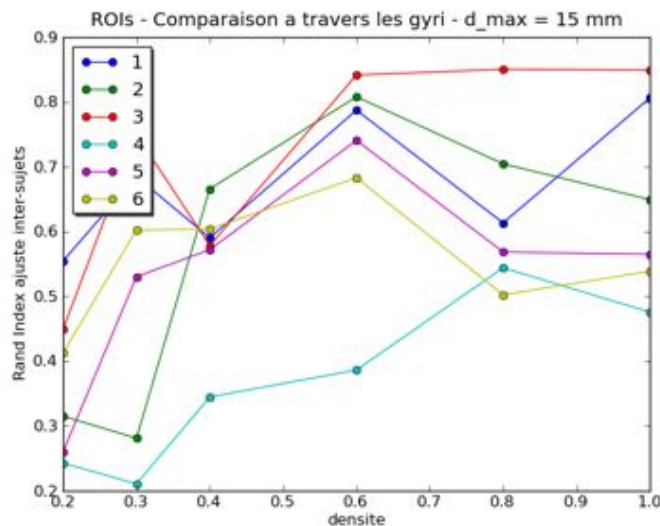


FIG. 7.12 – Sur 10 sujets. Comportement de la méthode M_{ROIs} sur les 6 gyri de la base simulée $B_{T_{density, \gamma=0.9}}$ ($d_{max} = 15mm$) pour différentes valeurs de densité de connexion.

Comportement des méthodes en fonction de la variabilité inter-individuelle La figure 7.14 présente le même type de graphique que celui de la figure 7.10 pour différents niveaux de variabilité inter-individuelle, allant d'un décalage maximum de 15 mm à 40 mm ($d_{max} \in [15, 20, 30, 40]$), et ceci pour 20 sujets. Ces graphiques sont assez

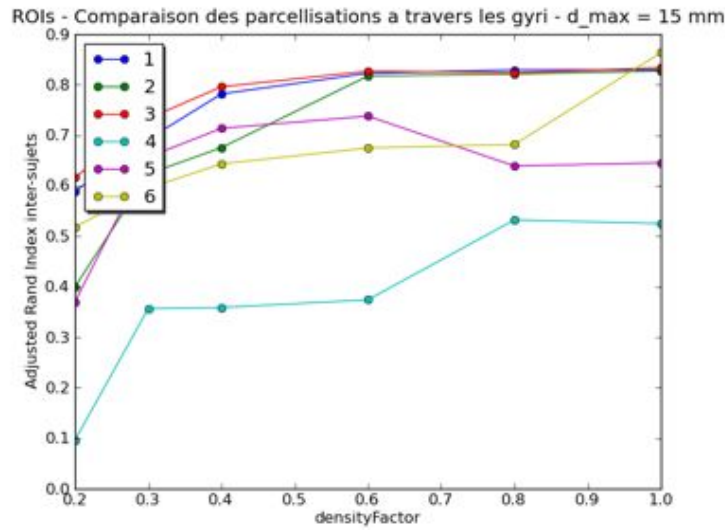


FIG. 7.13 – Sur 20 sujets. Comportement de la méthode M_{ROIs} sur les 6 gyri de la base simulée $B_{T_{density}, \gamma=0.9}(d_{max} = 15mm)$ pour différentes valeurs de densité de connexion.

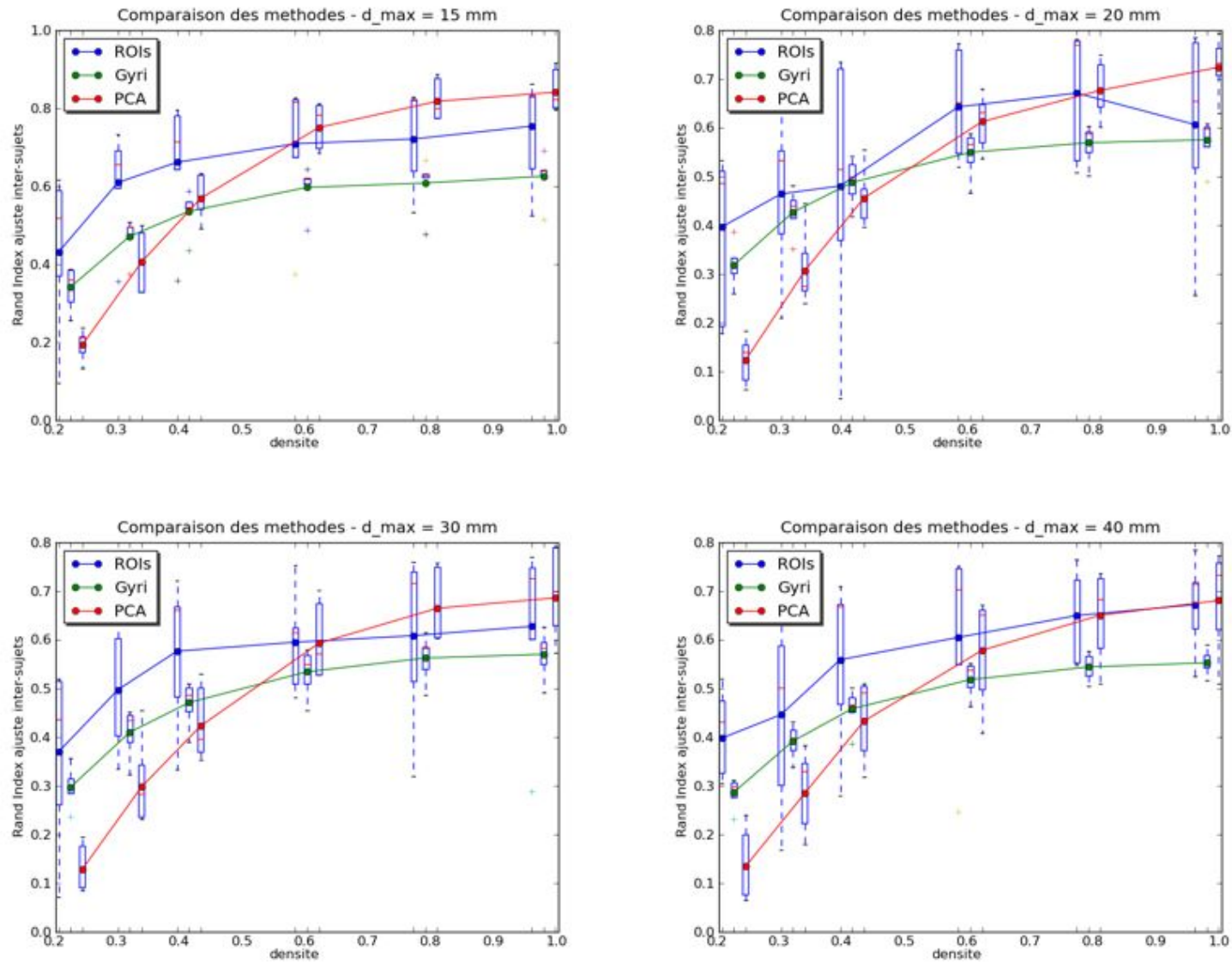
durs à interpréter. Ceci est peut-être dû au fait que la modélisation de la variabilité inter-individuelle dans les données simulées n'est pas totalement maîtrisée. En effet, le paramètre d_{max} dont elle dépend contrôle, pour chaque parcelle d'un sujet fils, le rayon du disque géodésique centré autour du centroïde de la parcelle du sujet modèle dans lequel le centroïde de la parcelle du sujet dérivé peut se trouver. L'attribution de la position du nouveau centroïde est faite de façon aléatoire. Une autre solution aurait été de la tirer d'une distribution gaussienne. Dans tous les cas, il faudrait augmenter le nombre de sujets.

Nous pouvons cependant constater que :

- La méthode M_{Gyri} ne présente pas de changement de comportement notable avec l'augmentation de la variabilité inter-individuelle, son Rand index ajusté restant toujours inférieur à M_{ROIs} .
- Le comportement des méthodes les unes par rapport aux autres reste cohérent, et notamment leur comportement quand la densité de connexion est faible.

Nous allons maintenant appliquer quelques unes de ces méthodes dans le cas des données réelles pour parcelliser la région englobant le SMA et le pré-SMA déjà bien étudiée dans la littérature.

FIG. 7.14 – Sur 20 sujets avec différents niveaux de variabilité inter-individuelle entre les parcellisations. Comparaison des méthodes inter-sujets sur la base simulée $B_{T_{density}, \gamma=0.9}$ ($d_{max} = 15\text{mm}$) pour différentes valeurs de densité de connexion.



7.2.4 Données réelles : région englobant le SMA et le pré-SMA

Nous allons ici présenter le comportement des différentes méthodes précédentes pour parcelliser la région $R_{SMA \cup \text{pré-SMA}}$ englobant le SMA et le pré-SMA, dont la définition et la construction sont décrites dans la partie 2.3.2.3.

Comme nous l'avons vu dans la partie 1.2.3.2, cette région a été divisée en 2 ou 3 régions à partir de données de connectivité anatomique. Un article (Crippa et al. 2010) identifie aussi deux régions avec une zone transitoire entre les deux.

Les approches précédentes utilisent soit une approche aveugle, soit une approche par régions d'intérêt définies à partir d'a priori sans tenir compte des données de connectivité en présence.

Rappelons que les méthodes de parcellisation que nous allons utiliser ici se situent dans une approche intermédiaire : la réduction des profils de connectivité se fait à partir des données de connectivité elles-mêmes sans utiliser d'a priori.

Données d'entrée Nous étudions ici la deuxième base de données constituée des 10 sujets de la base NMR. Les prétraitements effectués sont résumés dans la partie 2.4.2. Les données de **connectivité anatomique** sont issues d'un **tracking probabiliste** effectué sur les données HARDI de la base. La surface corticale de chaque sujet est issue d'une analyse Free-surfer fournissant des maillages du cortex avec une correspondance point à point à travers les sujets. Pour chaque sujet, la matrice de connectivité des points de la région d'intérêt est calculée suivant la méthode décrite à la partie 3.3.

Connectivité de $R_{SMA \cup \text{pré-SMA}}$ La figure 5.20 de la partie 5.2.5 sur la réduction des profils de connectivité nous permet de constater que les points de $R_{SMA \cup \text{pré-SMA}}$ diffèrent par leurs profils de connectivité vers les régions identifiées dans cette étude. Dans cette partie, nous allons présenter les résultats d'un clustering de cette région à partir des données de connectivité dans différents cas :

1. avec les profils réduits par l'identification automatique de régions d'intérêt : en inter-sujets : M_{ROIs}^{inter} (cf. Fig. 7.15, et 7.16),
2. avec les profils réduits par une analyse en composantes principales en intra-sujet : M_{PCA}^{intra} (cf. Fig. 7.17, 7.18, et 7.19).

7.2.4.1 Méthode M_{ROIs}^{inter}

La figure 7.15 présente les résultats du clustering inter-sujets à partir de l'identification de régions d'intérêt (méthode M_{ROIs}^{inter}) pour un nombre de clusters k de 2, 3 et 5, pour les 4 premiers sujets de la base. La largeur de silhouette moyenne des données est maximisée pour 5 clusters (cf. partie 6.2.2.2 pour une définition de ce critère). La ligne verticale passant par la commissure antérieure dans le repère de Talairach est représentée en rouge. D'après l'étude de Picard et Strick (1996) effectuée sur le macaque, cette ligne sépare cette région en deux sous-régions que sont le SMA et le pré-SMA. Nous constatons que :

- pour $k = 2$, cette ligne coïncide avec la frontière entre les deux clusters pour deux des quatre sujets représentés (sujet 2 et 3),
- pour $k = 3$, les frontières entre les clusters sont plus proches de cette ligne pour les quatre sujets,
- pour $k = 5$, le nombre de clusters jugé optimal i.e. maximisant la largeur de silhouette moyenne, il n’y a pas de grande différence aux niveaux des frontières des clusters proches de cette ligne.

La figure 7.16 montre la parcellisation obtenue pour le premier sujet ainsi que les profils de connectivité réduits des 3 parcelles médiales moyennés à travers tous les sujets.

Ainsi nous constatons que :

- La parcelle bleu foncé antérieure a des connexions avec le cortex frontal, nous pouvons l’associer au pré-SMA ;
- La parcelle vert clair a des connexions avec le cortex moteur et un nombre plus important de connexions corticospinales, nous pouvons l’associer au SMA,
- La parcelle bleu clair a un profil intermédiaire aux deux précédentes, avec un nombre plus important de connexions ventrales (la parcelle elle-même s’étendant ventralement).

Ceci est cohérent avec Crippa et al. (2010) qui met en évidence l’existence de deux régions distinctes séparées par une région de transition.

D’après les profils de connectivité, nous avons une idée des régions du cortex dont les connexions différencient les points de la région que l’on parcellise. Nous allons voir dans la suite les régions du cortex expliquant le mieux la variance des données contenues dans les profils de connectivité en effectuant une analyse en composantes principales sur les profils non réduits.

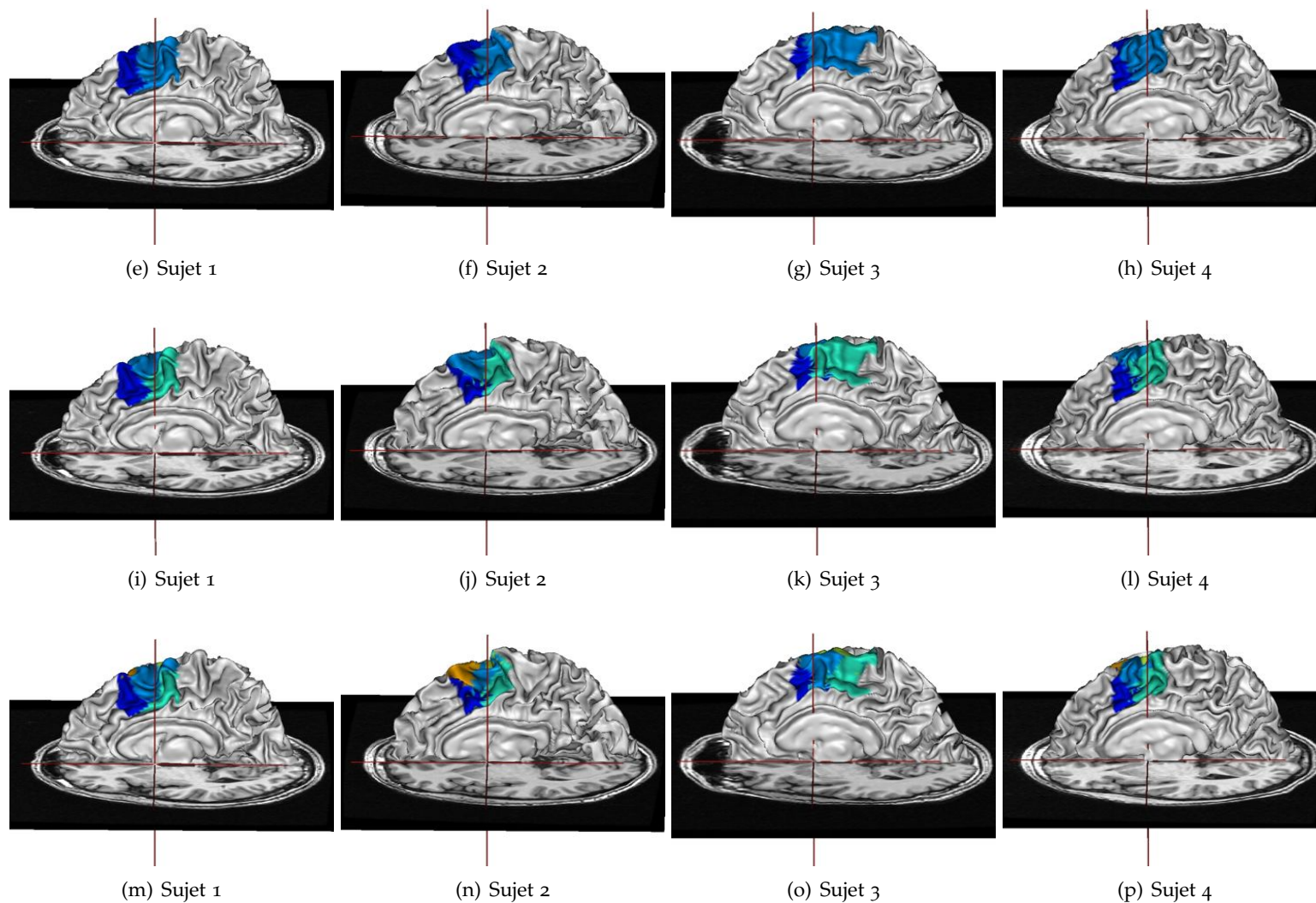
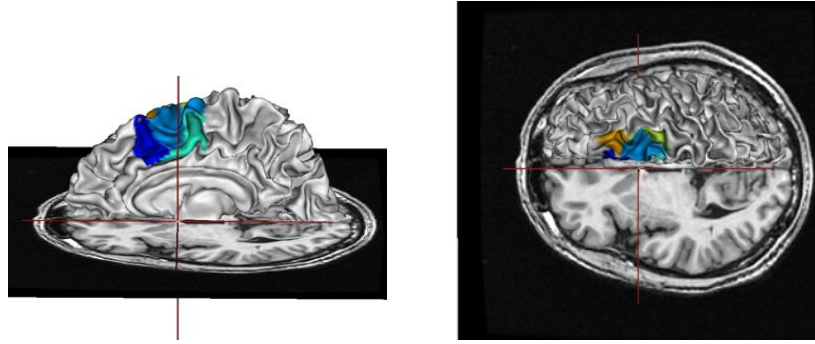
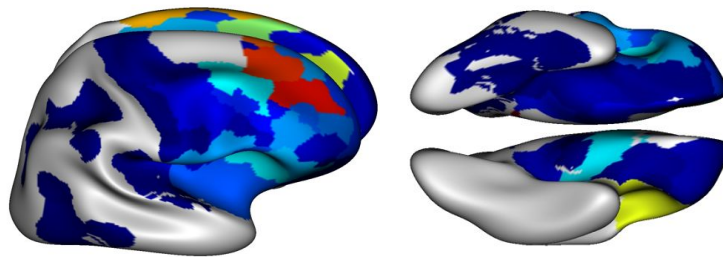


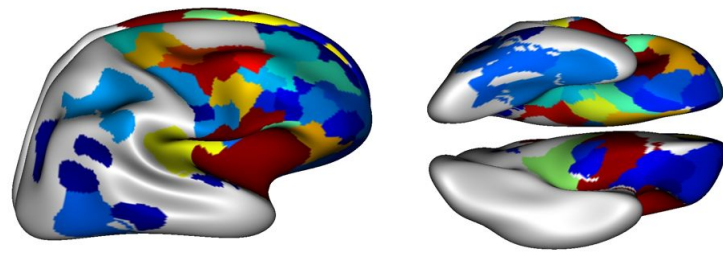
FIG. 7.15 – Parcellisations inter-sujets de la région englobant le SMA et le pré-SMA, rangées en fonction du nombre de clusters pris pour le clustering. Première ligne : 2 clusters, 2ème ligne : 3 clusters, 3ème ligne : 5 clusters (nombre de clusters optimal d'après l'avg silhouette width). La ligne verticale représente la ligne VAC qui est sensée séparer le CMA du pré-SMA.



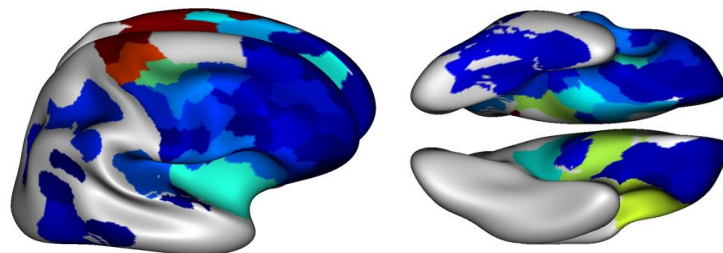
(a) Les parcelles obtenues pour le premier sujet



(b) Parcelle bleu foncé



(d) Parcelle bleu clair



(f) Parcelle vert clair

FIG. 7.16 – (a) : Parcellisation inter-sujets de la région englobant le SMA et le pré-SMA. (b-f) : Profils de connectivité des 3 parcelles médiales vers les bassins trouvés (somme des connexions pour l'ensemble des sujets, avec une échelle couleur croissante du bleu au bordeaux en fonction du pourcentage des connexions de chaque bassin dans les connexions de la parcelle considérée).

7.2.4.2 Méthode M_{PCA}^{intra}

J'ai effectué un clustering de la région englobant le SMA et le pré-SMA à partir des données de connectivité de l'hémisphère droit seulement (en comptant le corps calleux). Les tracts n'ont pas été seuillés en fonction de leur longueur.

ACP La figure 7.17 présente les premières composantes principales trouvées après avoir effectué une ACP sur ces données. Nous constatons que :

1. seules apparaissent dans l'analyse en composantes principales les régions connectées proches spatialement de $R_{SMA \cup \text{pré-SMA}}$,
2. les connexions inter-hémisphériques sont représentées dans les premières composantes (régions du corps calleux ici car les données de connectivité sont celles de l'hémisphère droit seulement).

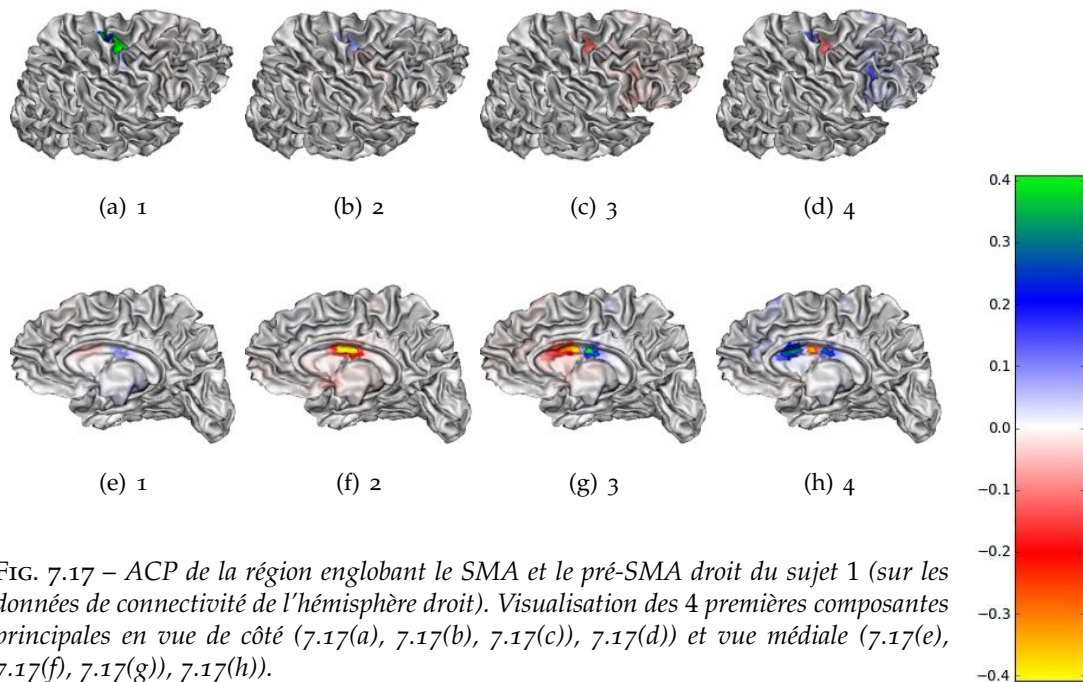


FIG. 7.17 – ACP de la région englobant le SMA et le pré-SMA droit du sujet 1 (sur les données de connectivité de l'hémisphère droit). Visualisation des 4 premières composantes principales en vue de côté (7.17(a), 7.17(b), 7.17(c)), 7.17(d)) et vue médiale (7.17(e), 7.17(f), 7.17(g)), 7.17(h)).

Parallèle avec l'étude d'Anwander et al. (2007) Le deuxième point est à mettre en lien avec l'étude d'Anwander et al. (2007) qui parcellisait simultanément les $R_{SMA \cup \text{pré-SMA}}$ des deux hémisphères, et qui trouvait au final 5 régions dont une apparentée au SMA qui regroupait des voxels des deux hémisphères, et 4 autres différenciées entre les hémisphères, apparentées aux cortex moteur et au pré-SMA de chacun des hémisphères. Le fait que le clustering regroupe des voxels des deux hémisphères dans cette région apparentée au SMA suggère, conformément à la PCA que nous avons fait, que les connexions inter-hémisphériques ont joué un rôle important dans le clustering. En effet, ce sont les seules connexions communes qui existent entre les voxels des deux hémisphères.

La figure 7.18 présente les scores des points de la région $R_{SMA \cup \text{pré-SMA}}$ suivant chacune des composantes principales précédentes. Cette visualisation permet ainsi d'avoir une idée du pouvoir de séparation de chacune des composantes sur la région.

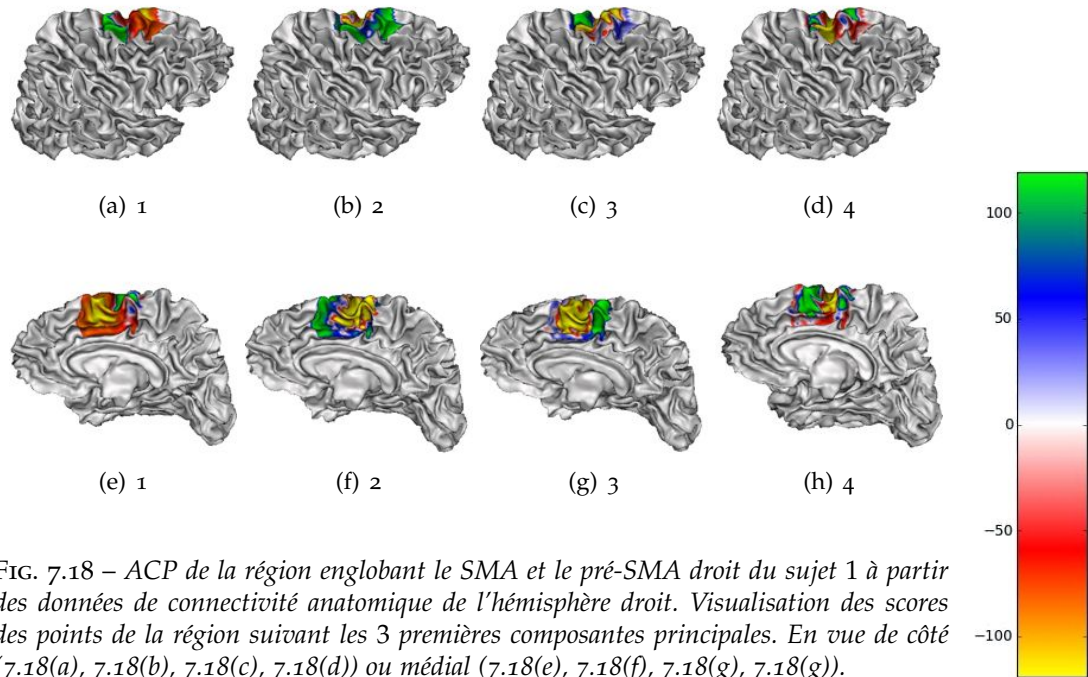


FIG. 7.18 – ACP de la région englobant le SMA et le pré-SMA droit du sujet 1 à partir des données de connectivité anatomique de l'hémisphère droit. Visualisation des scores des points de la région suivant les 3 premières composantes principales. En vue de côté (7.18(a), 7.18(b), 7.18(c), 7.18(d)) ou médial (7.18(e), 7.18(f), 7.18(g), 7.18(h)).

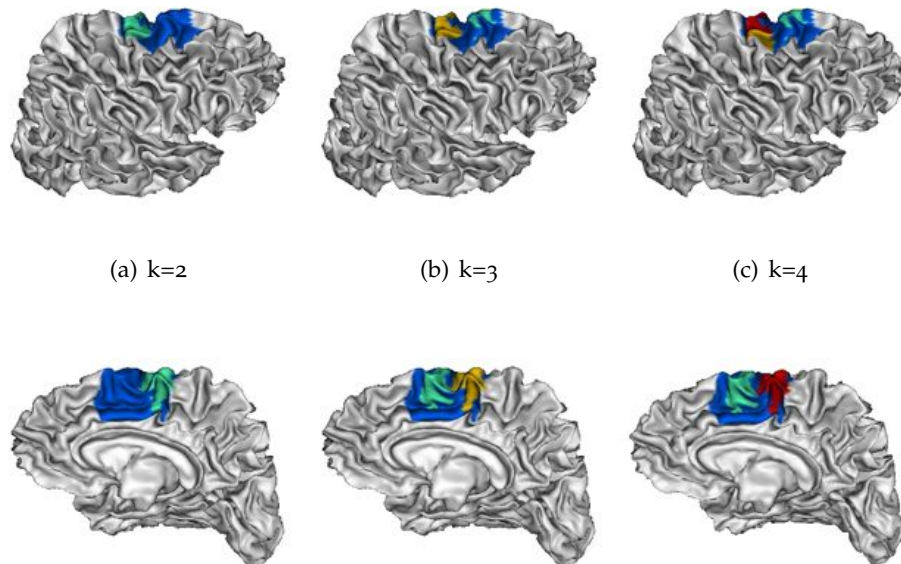


FIG. 7.19 – M_{PCA}^{intra} : Clustering après PCA de la région englobant le SMA et le pré-SMA droit du sujet 1 à partir des données de connectivité anatomique de l'hémisphère droit. Visualisation des parcellisations obtenues pour un nombre de clusters de 2 à 4 (7.19(a), 7.19(b), 7.19(c)).

Clustering sur les profils réduits La figure 7.19 montre les parcellisations obtenues en effectuant un k-médoides sur les profils réduits après analyse en composantes principales. Ainsi, nous constatons que pour un nombre de clusters égal à deux, seule la première composante de la PCA joue un rôle : la parcellisation divise la région $\mathbf{R}_{\text{SMAUp\acute{e}-SMA}}$ conformément à la carte des scores des points suivant cette première composante.

Les 3 clusters médians trouvés (dans le cas $k = 3$ ou $k = 4$) sont cohérents avec ceux trouvés dans l'approche $M_{\text{ROIs}}^{\text{inter}}$. Ce n'est pas le cas cependant des clusters trouvés pour $k = 2$.

7.2.4.3 Discussion

L'analyse en composantes principales permet de trouver les régions qui expliquent le mieux la variance des données et ainsi d'identifier la contribution de chacune de ces régions dans le clustering final à travers la visualisation de la carte spatiale des scores des points de la région suivant les différentes composantes.

Jbabdi et al. (2009), dans un cadre bayésien, a proposé un moyen de quantifier la contribution des différentes régions utilisées dans le clustering. Il utilise une segmentation a priori du cortex pour réduire les profils de connectivité, et il étudie donc la contribution des différentes régions de cette segmentation dans le clustering.

A la différence de Jbabdi et al. (2009), la méthode présentée ici (très simple puisque reposant sur une simple analyse en composantes principales) permet de quantifier la contribution des régions différenciant les points de $\mathbf{R}_{\text{SMAUp\acute{e}-SMA}}$ tout en les identifiant.

Ainsi, nous nous rendons compte ici que ce sont les régions les plus proches qui comptent le plus si aucun seuil par rapport à la longueur des connexions n'est effectué.

7.3 CONCLUSION

La parcellisation automatique de la surface corticale d'un groupe de sujets simultanément à partir des données de connectivité représente un objectif complexe et ambitieux. Dans la littérature, les études utilisent une segmentation a priori du cortex pour réduire la dimension des données et les placer dans une base commune au groupe de sujets considéré.

Dans la première partie de ce chapitre, nous avons proposé une démarche alternative à l'utilisation d'a priori en appliquant une méthode de réduction de dimension (présentée dans le Chapitre 5) qui tient compte de l'organisation spatiale des connexions et qui repose sur une mise en correspondance des maillages des différents sujets. Le clustering est restreint à une région corticale, un gyrus par exemple et consiste en deux approches : une approche dans laquelle les profils de connectivité de l'ensemble des sujets sont concaténés, et une approche dans laquelle ils sont moyennés. Nous avons appliqué cette méthode pour parcelliser le gyrus post-central. Les résultats obtenus sont cohérents entre les deux approches.

Dans la deuxième partie de ce chapitre, nous avons, dans le cas des données simulées, comparé la démarche précédente avec d'autres méthodes reposant sur des réductions de dimension des données différentes. Les jeux de données simulées ont ainsi permis d'évaluer le comportement des méthodes en fonction de la densité des connexions ainsi que de la variabilité inter-individuelle des parcellisations. Ainsi, il semblerait que les méthodes utilisant l'organisation spatiale des connexions se comportent mieux lorsque la densité des connexions est faible. Cette étude suggère de plus que si l'on utilise une segmentation a priori du cortex pour réduire les profils de connectivité, celle-ci ne doit pas être trop grossière. Il arrive en effet que la segmentation en gyri ne puisse pas suffire à séparer des parcelles dont les connexions tombent dans le même gyrus. Dans ce cas, l'approche par identification automatique de régions d'intérêt ou celle avec analyse en composantes principales s'avèrent être plus pertinentes.

Ces méthodes ont ensuite été appliquées à la parcellisation de la région englobant le SMA et le pré-SMA. L'analyse en composantes principales fournit alors un moyen d'identifier les régions expliquant le mieux la variance des données et indirectement celles qui comptent pour le clustering. Les résultats obtenus sont cohérents avec la littérature mais pas non plus satisfaisants. En effet, dans le cas des données réelles, la prépondérance des tracts courts par rapport aux tracts longue distance pose problème. Lorsqu'aucun seuil sur la longueur des connexions n'est appliqué, ce sont en effet les tracts courts qui jouent un rôle prépondérant dans le clustering. Ré-équilibrer la balance en faveur des tracts longs n'est pas évident, mais ce travail mérite d'être effectué. Une solution serait d'effectuer différents clusterings séparément pour différentes longueurs de tracts.

CONCLUSION GÉNÉRALE

LA parcellisation du cortex en aires fonctionnelles peut s'effectuer de diverses manières : il existe en effet de multiples techniques ainsi que différentes échelles spatiales où se positionner. Le sujet de cette thèse était la parcellisation corticale à partir des données de connectivité anatomique reconstruites par IRM de diffusion.

Dans ce contexte, une analyse de groupe semble être une bonne façon de s'affranchir ou du moins d'être moins sensible aux artéfacts et au caractère approximatif des faisceaux de neurones ou tracts reconstruits par tractographie. Cependant, la variabilité inter-individuelle est importante, que ce soit au niveau de l'anatomie sulco-gyrale, au niveau de l'activité cérébrale, ou au niveau des connexions de la substance blanche. Dans cette conclusion nous allons résumer les différents outils et méthodes que nous avons développés pour étudier et prendre en compte cette variabilité, dans le but de parcelliser le cortex à partir des connexions anatomiques corticales.

OUTILS ET MÉTHODES : DISCUSSION ET PERSPECTIVES

Construction et visualisation de connectomes anatomiques corticaux

Dans cette thèse, nous avons adopté une approche surfacique qui permet d'utiliser un système de coordonnées s'affranchissant d'une partie de la variabilité sulco-gyrale. Un recalage prenant en compte la morphologie du cortex et sa variabilité inter-individuelle est en effet effectué. Travailler sur la surface permet de plus de réduire la dimension des données et est cohérent avec la théorie des colonnes corticales.

Dans ce contexte, nous avons développé des outils pour construire et visualiser le connectome structurel surfacique qui permettent :

1. une **parcellisation du cortex avec des contraintes sulco-gyrales**, permettant ainsi de définir **des régions se retrouvant à travers les sujets**, en s'appuyant sur un système de coordonnées surfaciques,
2. le **calcul de la matrice de connectivité corticale**,
3. une **visualisation originale** du connectome structurel surfacique : avec un nouveau moyen de visualiser **la répartition spatiale** des connexions sur le cerveau entier.

Ces outils restent à compléter, notamment avec l'ajout des noyaux gris dans ce connectome. Le travail de Linda Marrackchi-Kacem (Marrackchi-Kacem et al. 2010) est en ceci complémentaire à celui effectué pendant

ma thèse. Il serait très intéressant d'effectuer une analyse du connectome similaire à celle d'Hagmann (Hagmann et al. 2008) incorporant les noyaux.

Simulation de connectomes avec contrôle de la variabilité

Nous avons proposé une méthode pour simuler des connectomes pour un ensemble d'individus, en modélisant la variabilité inter-individuelle et en contrôlant la densité des connexions. Les données simulées se construisent en deux étapes : (i) la création d'un sujet modèle et (ii) la création d'une base de sujets dérivés.

Dans cette thèse, nous avons testé différentes méthodes de parcellisation du cortex sur ce type de données simulées en simplifiant beaucoup les caractéristiques des connectomes ainsi construits (degrés des noeuds identiques, répartition uniforme des poids des connexions). Une suite logique de ce travail serait de se rapprocher plus des données réelles. Il serait par exemple possible d'introduire un biais en faveur des tracts courts, biais présent dans les données réelles issues de la plupart des algorithmes de tractographie. De plus, la surface corticale a été modélisée par une sphère pour simplifier, mais une extension possible serait de prendre les surfaces corticales réelles extraites à partir des IRM anatomiques. Enfin, la simulation au niveau du groupe peut aussi être améliorée.

Réduction inter-sujets des profils de connectivité

Lorsque l'on veut comparer, pour deux sujets différents, deux profils de connectivité entre eux, l'un des principaux enjeux est de gérer la variabilité interindividuelle des connexions des sujets.

Une solution est de regrouper les profils de connectivité à l'aide des régions d'une segmentation a priori du cortex, une segmentation en gyri par exemple. Nous avons dans cette thèse exploré deux méthodes de réduction des données de connectivité s'affranchissant de l'utilisation d'une telle segmentation a priori.

Identification automatique de régions d'intérêt

Pour une région corticale donnée R , nous avons proposé une méthode automatique permettant d'identifier un ensemble de régions d'intérêt connectées à R pour un sujet ainsi qu'au niveau d'un groupe de sujets. Cette méthode utilise les données de connectivité elles-mêmes, et s'effectue à partir d'une simple moyenne des données à travers les points de la région R . L'aspect spatial des connexions est pris en compte. Dans l'approche au niveau du groupe, le moyennage, effectué dans un espace commun à l'aide d'un système de coordonnées sphériques surfaciques permet de s'affranchir d'une partie de la variabilité inter-individuelle des données.

Cette méthode permet ainsi de résumer les connexions d'une région et simplifie leur visualisation. Dans le cas d'une région englobant le SMA et le pré-SMA, il est ainsi possible de visualiser d'un seul coup d'oeil les connexions de cette région. La décomposition effectuée par la réduction de dimension des données permet d'étudier les connexions de cette région

vers un même ensemble de régions pour les différents sujets, ce qui facilite les comparaisons inter-individuelles.

L'identification de ces régions d'intérêt fournit aussi une segmentation plus fine que les gyri pour réduire les profils de connectivité (dans la perspective d'effectuer un clustering sur ces profils). Nous proposons ainsi une approche intermédiaire entre l'approche aveugle, qui utilise les informations des connexions vers le cerveau entier et l'approche par régions d'intérêt définies de manière a priori.

Analyse en composantes principales

Nous avons étudié une autre technique de réduction de dimension classique : l'analyse en composantes principales ou PCA. Cette technique ne tient pas compte de l'information spatiale des connexions, mais l'ensemble des données de connectivité est utilisé : à partir du nuage de points formé par tous les points d'une région R dans l'espace de leur profil de connectivité cortical complet, un sous-espace maximisant la variance de ce nuage est trouvé. Dans le cas de la région $R_{SMA \cup \text{pré-SMA}}$ englobant le SMA et le pré-SMA, cette méthode identifie ainsi les connexions de la région considérée comme les maximas et minimas des axes (ou composantes) principaux de ce sous-espace. De plus, la visualisation des scores des points de la région suivant les différentes composantes est très informatif : ils varient spatialement en formant des zones homogènes et cohérentes.

Cette méthode fournit aussi la contribution des connexions (et des régions associées) dans la variance des données.

Discussion

Dans cette thèse, lorsque nous avons réduit les profils de connectivité par PCA sur la région englobant le SMA et le pré-SMA, les connexions qui expliquaient le mieux la variance des données étaient principalement les connexions courtes distance. Un travail futur intéressant serait **d'appliquer la méthode en filtrant les tracts suivant différentes longueurs de connexions séparément.**

Il serait de plus intéressant de combiner cette méthode avec la première méthode de réduction de dimension par identification de régions d'intérêt en effectuant une analyse en composantes principales sur les profils réduits vers ces régions. Ceci permettrait ainsi d'évaluer leurs contributions dans la variance des données.

Sur les données simulées, nous avons testé le comportement de ces deux méthodes pour différentes configurations possibles : différentes variabilité inter-sujets et différentes densités de connexions. La méthode par identification de régions d'intérêt est intéressante lorsque le nombre de connexions est faible, car l'information spatiale et le moyennage à travers les sujets la rend plus robuste que l'analyse en composantes principales. Cependant, la méthode par ROIs reste moins stable que la PCA lorsque la densité des connexions augmente. Durant cette thèse, nous avons travaillé sur une base de 10 sujets seulement. Il serait très intéressant **d'appliquer ces méthodes sur une plus grande base de sujets.** En effet, les données simulées ont l'air de montrer que la stabilité de la méthode avec ROIs augmente avec le nombre de sujets.

Parcellisation itérative intra-sujet du cortex entier

Dans cette partie nous avons proposé une méthode itérative ayant pour objectif de parcelliser le cortex entier en fonction des données de connectivité anatomique avec une double réduction de dimension : (i) une division du domaine à parcelliser, (ii) une réduction des profils de connectivité des points des sous-domaines précédemment définis par identification automatique de régions d'intérêt (cf. 7.3). Nous avons ensuite effectué un clustering sur ces données réduites à l'aide d'une version robuste du k-moyennes, le k-médoïdes, qui dispose d'une procédure pour déterminer le nombre optimal de clusters à choisir.

Cet objectif est très ambitieux car il s'interdit d'utiliser une segmentation a priori du cortex pour réduire les données. Il est de plus intra-sujet et le but est de parcelliser le cortex entier. Nous avons donc été conduits à formuler des hypothèses pour simplifier le problème, en ne prenant en compte que les tracts longs, ainsi qu'en fixant l'aire minimale d'une parcelle.

Les résultats sur 4 sujets sont plutôt encourageants mais ils suggèrent d'effectuer une analyse inter-sujets. En effet, sur une zone autour de la région de Broca, les différentes parcelles trouvées sont relativement bien séparées et les résultats sont relativement peu sensibles à l'initialisation. Cependant l'appariement des parcelles entre les sujets est difficile sans système de coordonnées surfaciques.

Discussion et perspectives

Une fois les parcellisations individuelles obtenues, un travail intéressant serait d'apparier automatiquement les parcelles à travers les sujets. Une solution serait d'**effectuer un clustering inter-sujets après le clustering intra-sujet**. Pour ce faire, il faut avoir un espace commun à travers les sujets pour comparer les parcelles et leur profil de connectivité. Cet espace peut nous être fourni par les méthodes de réduction de dimension inter-sujets que nous avons présentées plus haut.

Effectuer cette méthode itérative sur l'ensemble des sujets à la fois est plus compliqué, car le domaine à parcelliser de chaque sujet va évoluer différemment au cours des itérations. En revanche, à l'aide d'un système de coordonnées surfaciques, il est possible de calculer le connectome moyen d'un groupe de sujets. Cette méthode itérative pourrait alors être appliquée pour **parcelliser le cortex d'un sujet moyen construit à partir d'un groupe de sujets**. Une autre solution est d'effectuer une analyse inter-sujets mais sur une zone restreinte du cortex, ce que j'ai proposé par la suite.

Parcellisation de groupe du cortex

Parcellisation avec l'identification de régions d'intérêt

Dans cette partie nous proposons une procédure pour parcelliser le cortex d'un groupe de sujets simultanément, à partir des données de connectivité, à l'aide : (i) d'une mise en correspondance des sujets à travers un système de coordonnées surfaciques (Freesurfer), (ii) d'une réduction

de dimension des profils de connectivité effectuée au niveau du groupe, sans segmentation a priori mais à partir des données de connectivité elles-mêmes, et tenant compte de l'organisation spatiale des connexions, tout en étant peu coûteuse en calcul ; (iii) d'un clustering des profils réduits avec une version robuste du k-moyennes, le k-médoïdes, accompagné d'un critère pour choisir le nombre de clusters approprié.

Cette méthode appliquée à la parcellisation du gyrus post-central donne des parcelles cohérentes spatialement, homogènes et bien séparées les unes des autres en termes de profils de connectivité, avec correspondance entre les sujets immédiate. Certaines parcelles spécifiques à un groupe de sujets, peuvent de plus être identifiées.

Comparaison de différentes méthodes de réduction des profils

Sur une base de données simulées, nous avons comparé les méthodes de réduction des données de connectivité ainsi que les résultats d'un clustering inter-sujets effectué à partir de ces données réduites. Nous avons ainsi montré l'intérêt des méthodes tenant compte de l'organisation spatiale des données quand la densité des connexions est faible (méthode par identification de régions d'intérêt ou utilisant une segmentation a priori en gyri par exemple). L'analyse en composantes principales présente de bonnes performances quand la densité des connexions dépasse un certain seuil mais est plus gourmande en temps et mémoire de calcul.

Discussion et perspectives

Dans cette thèse, nous avons testé différentes méthodes pour trouver un espace approprié décrivant et résumant les données de connectivité d'un groupe de sujets. Nous avons ainsi proposé des façons de définir les régions qui vont influencer le clustering, sans utiliser de segmentation a priori du cortex.

Le problème reste difficile et plutôt mal posé. En effet, pour arriver à parcelliser une zone du cortex, nous avons cherché à identifier des régions à partir desquelles l'on regroupe les profils de connectivité de cette zone. Une parcellisation préalable du cortex est nécessaire dans cette approche.

Nous avons soit utilisé un simple moyennage des connexions à travers les sujets, soit effectué une analyse en composantes principales inter-sujets pour identifier ces régions. Une alternative serait d'utiliser un algorithme d'appariement plus évolué représentant les régions de chaque sujet sous forme de graphe. De telles approches ont déjà été proposées dans le cadre de l'analyse de données IRMf (Operto (2008), Thirion et al. (2006)).

Une autre démarche serait de proposer une **procédure itérative : à partir d'une première parcellisation du cortex, en effectuer une seconde, puis itérer le processus jusqu'à ce qu'un critère de stabilité (à définir) soit atteint.**

Une prochaine étape serait de comparer différentes méthodes de clustering et non pas les méthodes de réduction de dimension et de mise en correspondance des profils de connectivité préalable au clustering.

Les parcellisations obtenues restent cependant difficile à valider, car il n'existe pas à ce jour de gold standard en termes de parcellisation du cortex comme nous l'avons vu dans le chapitre introductif 1.2. Dans un futur proche, un gold standard pourrait émerger grâce à une technique d'imagerie en lumière polarisée (Palm et al. 2010). En effet, c'est une technique optique post mortem qui donne l'orientation des fibres à une résolution de 100 microns. Ceci pourrait faire grandement avancer les méthodes de parcellisation du cortex à partir des données de connectivité anatomiques : en permettant de valider les données de tractographie ainsi qu'en fournissant une parcellisation gold standard du cortex à partir des fibres.

APPLICATIONS

Applications méthodologiques

A ce jour, il n'y a pas de parcellisation universelle du cortex à l'échelle mésoscopique et de plus, la reconstruction des fibres à partir de l'IRM de diffusion est encore approximative et reste à valider. De plus, la variabilité interindividuelle est à prendre en compte. Dans ce contexte, il est important d'avoir des outils pour valider les études effectuées sur un connectome cérébral surfacique construit à partir de ces données. Les méthodes de parcellisation que j'ai présentées fournissent des pistes pour identifier les noeuds ou éléments de base d'un tel connectome. Des données simulées sont alors nécessaires pour valider et définir le cadre d'utilisation des différentes méthodes. Dans ce cadre, une méthode de simulation d'un ensemble de connectomes avec une modélisation de la variabilité interindividuelle telle que je l'ai développée pendant ma thèse paraît être appropriée. Elle reste bien sûr à améliorer. Il serait par exemple possible de simuler différentes répartitions entre les longueurs de tracts, et de favoriser ainsi les tracts courts ou longs.

De plus, comme nous l'avons vu, il existe aujourd'hui de nombreux algorithmes de tractographie. La méthode par identification automatique de régions d'intérêt et la visualisation originale que j'ai développées pourraient ainsi servir à comparer ces différents algorithmes. La visualisation permet notamment d'appréhender la répartition spatiale des connexions, et pourrait servir pour comparer le biais des différentes méthodes en faveur des tracts courts d'un seul coup d'oeil sur un cerveau entier.

Applications cliniques

Il y a un lien avéré entre l'altération des connexions et les maladies cérébrales, telles que l'autisme, la schizophrénie, ou la maladie d'Alzheimer par exemple. Etudier le connectome cérébral humain et trouver une représentation adaptée à chaque maladie pourrait apporter beaucoup dans la compréhension de ces pathologies.

Parcelliser le cortex d'une base de sujets en étant capable d'identifier des parcelles spécifiques aux sujets pathologiques pourrait alors permettre de faire émerger des différences d'organisation et de répartition des noeuds du connectome. Les méthodes de parcellisation inter-sujets

que nous avons présentées pendant cette thèse fournissent des pistes pour réaliser cet objectif, sur une zone restreinte du cortex.

Avant ce genre d'application, la visualisation originale que j'ai développée pourrait dès à présent être ajoutée aux différents types de visualisation dont les neuro-radiologues disposent. Combinée à une tractographie, elle permettrait de leur éviter la définition de régions d'intérêt pour visualiser les tracts tout en fournissant une information sur la connectivité du cortex entier.

Annexes

REVUE DES PARCELLISATIONS CORTICALES À PARTIR DE LA CONNECTIVITÉ ANATOMIQUE

A

Nous présentons dans cette annexe un tableau récapitulant les différentes études parcellisant le cortex en fonction de la connectivité anatomique reconstruite par IRM de diffusion. Pour chacune des études, nous précisons ainsi :

- les voxels graines servant pour l’initialisation des algorithmes de tractographie,
- les régions cibles qui sont les régions à partir desquelles sera décrite la connectivité des régions graines,
- la méthode utilisée pour la tractographie,
- la mesure de connectivité utilisée,
- les données d’entrée du clustering,
- la méthode de clustering utilisée,
- le nombre de clusters et comment ce choix est effectué (nous employons le terme **manuel** lorsque celui-ci a été déterminé manuellement après une étape de réordonnement de la matrice de corrélation ; nous le considérons comme **fixé** lorsqu’il est choisi par l’utilisateur avant l’étape de clustering),
- les méthodes utilisées pour la validation des résultats obtenus.

Nous avons classé les études selon la région du cortex étudiée ainsi que selon leur caractère intra- ou inter-sujets.

continued from previous page

Etude	Profils		Tracto.	Mesure de connectivité	Clustering		Nombre de clusters	Validation
	Régions graines	Régions cibles			Données	Méthode		
Broca								
Anwander et al. (2007)	voxel de la région interface WM/GM $1 \times 1 \times 1 \text{ mm}^3$	idem Anwander et al. (2007) présenté au-dessus			Matrice de corrélation	k-means	3 (BA44, 45 et deep frontal operculum)	6 sujets, Comparaison avec cartes probabilistes des aires cytoarchitectoniques BA44 et 45
Klein et al. (2007) idem SMA/pré-SMA présenté au-dessus							2, fixé	
Insula								
Nanetti et al. (2009) idem SMA/pré-SMA présenté au-dessus							2, fixé	
Cortex cingulaire								
Beckmann et al. (2009)	masque de la région identique pour tous les sujet défini dans le MNI	voxels du reste du cerveau $5 \times 5 \times 5 \text{ mm}^3$	probabiliste, 1 000 part. par voxel graine	proba de connexion pondérée par la longueur	Matrice de corrélation	k-means avec plusieurs itérations	9, choisis d'après la consistance inter-individus	11 sujets, Comparaison avec cartes de localisation fonctionnelle construites après méta-analyse de 171 études avec des activations fonctionnelles dans la région
	idem	11 régions définies manuellement	probabiliste, 5 000 part./voxel graine	idem	pour chaque région cible : observation des voxels cingulaires connectés de manière reproductible à travers les sujets, et lien avec les 9 clusters obtenus par l'autre approche			
Cortex entier								
Perrin et al. (2008)	Masque de la matière blanche tenant compte des sillons	36 régions du cortex (gyri)	probabiliste	nombre de particules	connectivité des N voxels de chaque gyrus vers les 36 gyri, normalisée	k-means régularisé	-	3 sujets
INTER-SUJETS								
SMA/pré-SMA								
Jbabdi et al. (2009)	masque dans l'espace du MNI	48 ou 7 régions couvrant le cortex	probabiliste, 5 000 part./graine	nombre de particules puis log et ajout de bruit gaussien	Matrice de connectivité vers 7 ou 48 régions	hierarchial Dirichlet process mixture models	2 clusters principaux, automatique	19 sujets, simulation

continued on next page

continued from previous page

Etude	Profils		Tracto.	Mesure de connectivité	Clustering		Nombre de clusters	Validation
	Régions graines	Régions cibles			Données	Méthode		
Gyrus post-central								
Guevara et al. (2008)	région surfacique	36 gyri surfaciques	probabiliste	nombre de particules	idem Per-rin et al. (2008) mais surfacique	k-means régularisé ou non	6 ou 8, fixé	3 sujets, reproductibilité dans la localisation des clusters à travers les sujets

ÉLÉMENTS DE THÉORIE DES GRAPHS

B

Soit $G = (\mathcal{V}, \mathcal{E})$ un graphe constitué :
– d'un ensemble de V noeuds $\mathcal{V} = \{v_1, \dots, v_V\}$
– et de E relations \mathcal{E} entre les noeuds.
On peut définir :

La matrice d'adjacence de G de taille (V, V) , booléenne (graphe binaire) ou à valeur dans \mathbb{R}^+ (graphe pondéré) A avec :

$$a_{ij} : \begin{cases} w & \text{si } (v_i, v_j) \in \mathcal{E} \\ 0 & \text{sinon.} \end{cases}$$

avec pour un graphe binaire $w = 1$ et $w \in \mathbb{R}^+$ pour un graphe pondéré.

Le degré d'un noeud est le nombre de noeuds du graphe auquel il est connecté. Le degré de tous les noeuds résulte en la **distribution des degrés du graphe** (degree distribution).

La force (strength) d'un noeud : la somme des poids des relations du noeud (égal au degré pour le graphe binaire).

La longueur de chemin (path length) entre deux noeuds représente le nombre minimum de relations qu'il faut parcourir pour aller d'un noeud à l'autre. On utilise aussi comme mesure globale du graphe la longueur de chemin moyenne (mean ou **average path length**) sur l'ensemble des noeuds du graphe.

Le coefficient de clustering d'un noeud mesure la densité de connexions au sein des plus proches voisins du noeud. Ainsi un noeud aura un fort coefficient de clustering s'il est entouré de noeuds très inter-connectés les uns aux autres.

Références

BIBLIOGRAPHIE

- S. Achard et E. Bullmore. Efficiency and cost of economical brain functional networks. *PLoS computational biology*, 3(2), 2007. (Cité page 15.)
- Y. Alemán-Gómez, L. Melie-García, et P. Valdés-Hernandez. IBASPM : Toolbox for automatic parcellation of brain structures. Dans *12th Annual Meeting of the Organization for Human Brain Mapping, June*, pages 11–15, 2006. (Cité page 18.)
- DC Alexander, GJ Barker, et SR Arridge. Detection and modeling of non-gaussian apparent diffusion coefficient profiles in human brain data. *Magnetic Resonance in Medicine*, 48(2) :331–340, 2002. (Cité page 51.)
- K. Amunts et K. Zilles. Advances in cytoarchitectonic mapping of the human cerebral cortex. *Neuroimaging Clinics of North America*, 11(2) :151, 2001. ISSN 1052-5149. (Cité page 30.)
- J. Annese, A. Pitiot, ID Dinov, et AW Toga. A myelo-architectonic method for the structural classification of cortical areas. *NeuroImage*, 21(1) :15–26, 2004. ISSN 1053-8119. (Cité page 19.)
- Alfred Anwander, M. Tittgemeyer, D.Y. von Cramon, Angela D. Friederici, et Thomas R. Knösche. Connectivity-based parcellation of broca’s area. *Cerebral Cortex*, 17 :816–825, April 2007. (Cité pages 23, 24, 25, 26, 48, 168, 184 et 185.)
- P.J. Basser, J. Mattiello, D. LEBiHAN, et al. Estimation of the effective self-diffusion tensor from the NMR spin echo. *Journal of Magnetic Resonance-Series B*, 103(3) :247–254, 1994. ISSN 1064-1866. (Cité pages 43 et 52.)
- D.S. Bassett, E. Bullmore, B.A. Verchinski, V.S. Mattay, D.R. Weinberger, et A. Meyer-Lindenberg. Hierarchical organization of human cortical networks in health and schizophrenia. *Journal of Neuroscience*, 28(37) : 9239, 2008. (Cité page 14.)
- D.S. Bassett et E.T. Bullmore. Human brain networks in health and disease. *Current opinion in neurology*, 22(4) :340, 2009. (Cité page 9.)
- Matthias Beckmann, Heidi J ohansen Berg, et Matthew F. S. Rushworth. Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. *The Journal of Neuroscience*, 29(4) : 1175–1190, January 2009. (Cité pages 23, 24 et 185.)
- T. E. J. Behrens, Heidi Johansen-Berg, Saad Jbabdi, M. F. S. Rushworth, et M. W. Woolrich. Probabilistic diffusion tractography with multiple fibre orientations : What can we gain? *NeuroImage*, 34 :144–155, 2007. (Cité page 53.)

- T. E. J. Behrens, Heidi Johansen-Berg, M. W. Woolrich, Stephen M Smith, C. A. M. Wheeler-Kingshott, P. A. Boulby, G. J. Barker, Emma L. Sillery, K. Sheehan, Olga Ciccarelli, Alan J. Thompson, J. M. Brady, et Paul M. Matthews. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. *Nature Neuroscience*, 6(7) : 750–757, July 2003. (Cité page 23.)
- M.K. Belmonte, G. Allen, A. Beckel-Mitchener, L.M. Boulanger, R.A. Carper, et S.J. Webb. Autism and abnormal development of brain connectivity. *Journal of Neuroscience*, 24(42) :9228, 2004. (Cité page 14.)
- J.W. Bohland, H. Bokil, C.B. Allen, et P.P. Mitra. The brain atlas concordance problem : quantitative comparison of anatomical parcellations. 2009a. (Cité pages 139 et 159.)
- J.W. Bohland, C. Wu, H. Barbas, H. Bokil, M. Bota, H.C. Breiter, H.T. Cline, J.C. Doyle, P.J. Freed, R.J. Greenspan, et al. A proposal for a coordinated effort for the determination of brainwide neuroanatomical connectivity in model organisms at a mesoscopic scale. *PLoS Comput Biol*, 5(3) : e1000334, 2009b. (Cité pages 10 et 13.)
- S.L. Bressler et V. Menon. Large-scale brain networks in cognition : emerging methods and principles. *Trends in Cognitive Sciences*, 2010. (Cité pages 7 et 13.)
- K. Brodmann. *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues*. Barth Leipzig, 1909. (Cité page 19.)
- R.L. Buckner, J. Sepulcre, T. Talukdar, F.M. Krienen, H. Liu, T. Hedden, J.R. Andrews-Hanna, R.A. Sperling, et K.A. Johnson. Cortical hubs revealed by intrinsic functional connectivity : mapping, assessment of stability, and relation to Alzheimer’s disease. *Journal of Neuroscience*, 29(6) :1860, 2009. (Cité pages 14 et 15.)
- Ed Bullmore et Olaf Sporns. Complex brain networks : graph theoretical analysis of structural and functional systems. *Nature Rev Neurosci*, 10 (3) :186–98, March 2009. (Cité page 14.)
- ET Bullmore, S. Frangou, et RM Murray. The dysplastic net hypothesis : an integration of developmental and dysconnectivity theories of schizophrenia. *Schizophrenia research*, 28(2-3) :143–156, 1997. ISSN 0920-9964. (Cité page 14.)
- U. Bürgel, K. Amunts, L. Hoemke, H. Mohlberg, J.M. Gilsbach, et K. Zilles. White matter fiber tracts of the human brain : three-dimensional mapping at microscopic resolution, topography and intersubject variability. *Neuroimage*, 29(4) :1092–1105, 2006. (Cité pages 31 et 32.)
- A. Cachia, JF Mangin, D. Riviere, D. Papadopoulos-Orfanos, F. Kherif, I. Bloch, et J. Régis. A generic framework for parcellation of the cortical surface into gyri using geodesic voronoi diagrams. *Medical Image Analysis*, 7(4) :403–416, 2003. (Cité pages 19 et 65.)

- P.T. Callaghan. Principles of nuclear magnetic resonance microscopy. 1993. (Cité page 52.)
- PT Callaghan, D. MacGowan, KJ Packer, et FO Zelaya. High-resolution q-space imaging in porous structures. *Journal of magnetic resonance*, 90 (1) :177–182, 1990. ISSN 0022-2364. (Cité page 49.)
- Leila Cammoun, Xavier Gigandet, Patric Hagmann, Reto Meuli, et Jean-Philippe Thiran. Diffusion MR Image Segmentation : Towards Global Brain Connectivity Analysis. Dans *Proceedings of the 10th European Congress of Stereology and Image Analysis*, 2009. (Cité page 16.)
- Pascal Cathier et Jean-François Mangin. Registration of cortical connectivity matrices. *MMBIA2006*, 2006. New York, USA. (Cité pages 63 et 68.)
- P. Ciuciu, J.B. Poline, G. Marrelec, J. Idier, C. Pallier, et H. Benali. Unsupervised robust nonparametric estimation of the hemodynamic response function for any fMRI experiment. *Medical Imaging, IEEE Transactions on*, 22(10) :1235–1251, 2003. ISSN 0278-0062. (Cité page 30.)
- C. Clouchoux, O. Coulon, J.L. Anton, J.F. Mangin, et J. Régis. A new cortical surface parcellation model and its automatic implementation. volume LNCS 4141, pages 193–200, Copenhagen, Danemark, 2006. Springer Verlag. (Cité pages 19 et 21.)
- C. Clouchoux, O. Coulon, D. Riviere, A. Cachia, JF Mangin, et J. Régis. Anatomically constrained surface parameterization for cortical localization. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2005*, pages 344–351, 2005. (Cité pages 19 et 142.)
- Alexander L. Cohen, Damien A. Fair, Nico U. F. Dosenbach, Francis M. Miezin, Donna L. Dierker, David C. Van Essen, Bradley L. Schlaggar, et Steven E. Petersen. Defining functional areas in individual human brains using resting functional connectivity mri. *NeuroImage*, 41 :45–57, 2008. (Cité page 20.)
- A. Crippa, L. Cerliani, L. Nanetti, et J.B.T.M. Roerdink. Heuristics for connectivity-based brain parcellation of sma/pre-sma through force-directed graph layout. *NeuroImage*, 2010. (Cité pages 23, 25, 77, 164, 165 et 184.)
- Anders M. Dale, Bruce Fischl, et Martin I. Serenot. Cortical surface-based analysis. i. segmentation and surface reconstruction. *NeuroImage*, 9 : 179–194, 1999. (Cité pages 44 et 46.)
- M.T. De Schotten et al. Atlasing location, asymmetry and inter-subject variability of white matter tracts in the human brain with mr diffusion tractography. *NeuroImage*, 54 :49–59, 2011. (Cité page 32.)
- M. Descoteaux. *High angular resolution diffusion MRI : From local estimation to segmentation and tractography*. PhD thesis, University of Nice-Sophia Antipolis, Nice, France, 2008. (Cité page 56.)

- M. Descoteaux, E. Angelino, S. Fitzgibbons, et R. Deriche. Regularized, fast, and robust analytical q-ball imaging. *Magnetic Resonance in Medicine*, 58(3) :497–510, 2007. (Cité pages 44, 52, 55, 56 et 155.)
- Maxime Descoteaux, Rachid Deriche, Thomas R. Knösche, et Alfred Anwander. Deterministic and probabilistic tractography based on complex fibre orientation distributions. *IEEE Transactions on Medical Imaging*, 28 : 269–286, February 2009. (Cité page 56.)
- R.S. Desikan, F. Ségonne, B. Fischl, B.T. Quinn, B.C. Dickerson, D. Blacker, R.L. Buckner, A.M. Dale, R.P. Maguire, B.T. Hyman, et al. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Neuroimage*, 31(3) :968–980, 2006. (Cité pages 18, 19, 44, 46, 66 et 148.)
- P.E. Downing, Y. Jiang, M. Shuman, et N. Kanwisher. A cortical area selective for visual processing of the human body. *Science*, 293(5539) :2470, 2001. (Cité page 13.)
- Qiang Du, Vance Faber, et Max Gunzburger. Centroidal voronoi tessellations : Applications and algorithms. *Society for Industrial and Applied Mathematics Review*, 41(4) :637–676, 1999. (Cité page 112.)
- Jeff H. Duyn, Peter van Gelderen, Tie-Qiang Li, et Jacco A. de Zwart. High-field mri of brain cortical substructure based on signal phase. *PNAS*, 104 : 11796–11801, 2007. (Cité page 19.)
- C. Economo et G.N. Koskinas. *Die cytoarchitektonik der hirnrinde des erwachsenen menschen*. J. Springer, 1925. (Cité page 19.)
- S.B. Eickhoff, T. Paus, S. Caspers, M.H. Grosbras, A.C. Evans, K. Zilles, et K. Amunts. Assignment of functional activations to probabilistic cytoarchitectonic areas revisited. *Neuroimage*, 36(3) :511–521, 2007. ISSN 1053-8119. (Cité page 30.)
- Simon B. Eickhoff, Klaas E. Stephan, Hartmut Mohlberg, Christian Grefkes, Gereon R. Fink, Katrin Amunts, et Karl Zilles. A new spm toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage*, 25 :1325–1335, 2005. (Cité pages 30 et 31.)
- A. Einstein. Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der physik*, 322(8) :549–560, 1905. ISSN 1521-3889. (Cité page 41.)
- B. Fischl, M.I. Sereno, R.B.H. Tootell, et A.M. Dale. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, 8(4) :272–284, 1999a. (Cité pages 44, 46 et 48.)
- Bruce Fischl, Niranjini Rajendran, Evelina Busa, Jean Augustinack, Oliver Hinds, B.T. Thomas Yeo, H. Mohlberg, K. Amunts, et K. Zilles. Cortical folding patterns and predicting cytoarchitecture. *Cerebral Cortex*, 18 : 1973–1980, 2007. (Cité pages 30 et 31.)

- Bruce Fischl, Martin I. Serenot, et Anders M. Dale. Cortical surface-based analysis. ii. inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9 :195–207, 1999b. (Cité pages 44 et 46.)
- G. Flandin, F. Kherif, X. Pennec, D. Riviere, N. Ayache, et J.B. Poline. Parcellation of brain images with anatomical and functional constraints for fMRI data analysis. 2002. (Cité page 20.)
- G. Flandin, W. Penny, X. Pennec, N. Ayache, et J.B. Poline. A multisubject anatomo-functional parcellation of the brain. *Neuroimage*, 19 :1600, 2003. (Cité page 20.)
- Guillaume Flandin. *Utilisation d'informations géométriques pour l'analyse statistique des données d'IRM fonctionnelle*. PhD thesis, Université de Sophia Antipolis, 2004. (Cité page 65.)
- L.R. Frank. Anisotropy in high angular resolution diffusion-weighted mri. *Magnetic Resonance in Medicine*, 45(6) :935–939, 2001. (Cité page 52.)
- W.J. Freeman. *Mass action in the nervous system*. Academic Press New York, 1975. (Cité page 13.)
- J.M. Fuster. The Module : Crisis of a Paradigm. *Neuron*, 26(1) :51–53, 2000. ISSN 0896-6273. (Cité page 13.)
- J.M. Fuster. *Cortex and mind : Unifying cognition*. Oxford University Press, USA, 2005. ISBN 019530084X. (Cité page 13.)
- N. Geschwind. Disconnexion syndromes in animals and man. *Brain*, 88 (3) :585, 1965. ISSN 0006-8950. (Cité page 14.)
- X. Gigandet. *Global Brain Connectivity Analysis by Diffusion MR Tractography : Algorithms, Validation and Applications*. PhD thesis, École Polytechnique Fédérale de Lausanne, 2009. (Cité page 65.)
- G. Goldberg. Supplementary motor area structure and function : review and hypotheses. *Behavioral and Brain Sciences*, 8(04) :567–588, 1985. ISSN 1469-1825. (Cité page 24.)
- P.S. Goldman-Rakic. Topography of cognition : parallel distributed networks in primate association cortex. *Annual Review of Neuroscience*, 11 (1) :137–156, 1988. ISSN 0147-006X. (Cité page 13.)
- P. Golland, Y. Golland, et R. Malach. Detection of spatial activation patterns as unsupervised segmentation of fMRI data. *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2007*, pages 110–118, 2007. (Cité page 20.)
- Gaolang Gong, Yong He, Luis Concha, Catherine Lebel, Donald W. Gross, Alan C. Evans, et Christian Beaulieu. Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. *Cerebral Cortex*, June 2008. Advance access. (Cité pages 14, 15 et 18.)

- J.C. Gower. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53(3-4) :325, 1966. ISSN 0006-3444. (Cité page 100.)
- P. Guevara, M. Perrin, P. Cathier, Y. Cointepas, D. Riviere, C. Poupon, et J.F. Mangin. Connectivity-based parcellation of the cortical surface using q-ball imaging. pages 903–906, 2008. (Cité pages 23, 24, 147 et 186.)
- D.A. Gusnard et M.E. Raichle. Searching for a baseline : functional imaging and the resting human brain. *Nature Reviews Neuroscience*, 2(10) : 685–694, 2001. (Cité page 20.)
- P. Hagmann, M. Kuran, X. Gigandet, P. Thiran, V.J. Wedeen, R. Meuli, et J.P. Thiran. Mapping human whole-brain structural networks with diffusion MRI. *PLoS One*, 2(7) :597, 2007. ISSN 1932-6203. (Cité pages 12 et 18.)
- Patric Hagmann, Leila Cammoun, Xavier Gigandet, Reto Meuli, Christopher J. Honey, Van J. Wedeen, et Olaf Sporns. Mapping the structural core of human cerebral cortex. *PLOS Computational Biology*, 6(7) :1479–1493, July 2008. (Cité pages 12, 14, 15, 17, 18, 26, 61, 65, 66, 67, 70, 72, 81 et 174.)
- JA Hartigan et MA Wong. A k-means clustering algorithm. *Journal of the Royal Statistical Society C*, 28(1) :100–108, 1979. (Cité page 66.)
- Yong He et Alan C. Chen, Zhang J. and Evans. Small-world anatomical networks in the human brain revealed by cortical thickness from mri-based. *Cerebral Cortex*, 17 :2407–2419, October 2007. (Cité pages 15 et 26.)
- H. Hotelling. Analysis of a complex of statistical variables into principal components. *Journal of educational psychology*, 24(6) :417–441, 1933. ISSN 0022-0663. (Cité page 111.)
- L. Hubert et P. Arabie. Comparing partitions. *Journal of classification*, 2(1) : 193–218, 1985. (Cité pages 139 et 159.)
- S.A. Huettel, A.W. Song, et G. McCarthy. Functional magnetic resonance imaging. 2004. (Cité pages 40 et 41.)
- Y. Iturria-Medina, E.J. Canales-Rodriguez, et L. Melie-Garcia. Characterizing brain anatomical connections using diffusion weighted mri and graph theory. *NeuroImage*, 36 :645–660, 2007. (Cité page 18.)
- Y. Iturria-Medina, R.C. Sotero, E.J. Canales-Rodríguez, Y. Alemán-Gómez, et L. Melie-García. Studying the human brain anatomical network via diffusion-weighted MRI and Graph Theory. *Neuroimage*, 40(3) :1064–1076, 2008. ISSN 1053-8119. (Cité page 14.)
- E.D. Jarvis, O. Güntürkün, L. Bruce, A. Csillag, H. Karten, W. Kuenzel, L. Medina, G. Paxinos, D.J. Perkel, T. Shimizu, et al. Avian brains and a new understanding of vertebrate brain evolution. *Nature Reviews Neuroscience*, 6(2) :151–159, 2005. ISSN 1471-003X. (Cité page 14.)

- S. Jbabdi, P. Bellec, R. Toro, J. Daunizeau, M. Pélégriani-Issac, et H. Benali. Accurate anisotropic fast marching for diffusion-based geodesic tractography. *Journal of Biomedical Imaging*, 2008 :1–12, 2008. (Cité page 53.)
- Saad Jbabdi. A bayesian framework for global tractography. *NeuroImage*, 37 :116–129, 2007. (Cité page 53.)
- Saad Jbabdi, M. W. Woolrich, et T. E. J. Behrens. Multiple-subjects connectivity-based parcellation using hierarchical dirichlet process mixture models. *NeuroImage*, 44 :373–384, 2009. (Cité pages 23, 24, 147, 170 et 185.)
- Heidi Johansen-Berg, T. E. J. Behrens, M. D. Robson, M. F. S. Rushworth, J. M. Brady, et Stephen M. Smith. Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *PNAS*, 101 :13335–13340, 2004. (Cité pages 23, 24, 25, 27, 48, 112 et 184.)
- N. Kanwisher, J. McDermott, et M.M. Chun. The fusiform face area : a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17(11) :4302, 1997. (Cité page 13.)
- S. Karkar, S. Faisan, L. Thoraval, et JR Foucher. A multi-level parcellation approach for brain functional connectivity analysis. Dans *Engineering in Medicine and Biology Society*, 2009. *EMBC 2009. Annual International Conference of the IEEE*, pages 3497–3500. IEEE, 2009. (Cité page 20.)
- B. Kastler, D. Vetter, Z.Â. Patay, et P. Germain. *Comprendre l'IRM : manuel d'auto-apprentissage*. Elsevier Masson, 2006. ISBN 2294051106. (Cité page 41.)
- L. Kaufmann et P. J. Rousseeuw. *Finding groups in data : an introduction to cluster analysis*. Wiley Interscience, 1990. (Cité pages 137, 138 et 150.)
- J.H. Kim et V.H. Vu. Generating random regular graphs. Dans *Proceedings of the thirty-fifth annual ACM symposium on Theory of computing*, pages 213–222. ACM, 2003. ISBN 1581136749. (Cité page 86.)
- Johannes C. Klein, T. E. J. Behrens, M. D. Robson, Clare E. Mackay, Desmond J. Higham, et Heidi Johansen-Berg. Connectivity-based parcellation of human cortex using diffusion mri : Establishing reproducibility, validity and observer independence in ba 44/45 and sma/pre-sma. *NeuroImage*, 34 :204–211, 2007. (Cité pages 23, 24, 184 et 185.)
- D. Le Bihan. Looking into the functional architecture of the brain with diffusion MRI. *Nature Reviews Neuroscience*, 4(6) :469–480, 2003. ISSN 1471-003X. (Cité page 42.)
- D. Le Bihan, E. Breton, D. Lallemand, P. Grenier, E. Cabanis, et M. Laval-Jeantet. MR imaging of intravoxel incoherent motions : application to diffusion and perfusion in neurologic disorders. *Radiology*, 161(2) :401, 1986. ISSN 0033-8419. (Cité pages 41 et 43.)
- S. Lehericy, M. Ducros, A. Krainik, C. Francois, P.F. Van de Moortele, K. Ugurbil, et D.S. Kim. 3-d diffusion tensor axonal tracking shows

- distinct sma and pre-sma projections to the human striatum. *Cerebral Cortex*, 14(12) :1302–1309, 2004. (Cité page 48.)
- C. Lenglet. Geometric and variational methods for diffusion tensor mri processing. 2006. (Cité page 53.)
- A.R. Luria, H.L. Teuber, K.H. Pribram, et B. Haigh. *Higher cortical functions in man*. Basic Books New York, 1966. (Cité page 13.)
- J.F. Mangin. Entropy minimization for automatic correction of intensity nonuniformity. pages 162–169, 2000. (Cité page 45.)
- J.F. Mangin, V. Frouin, I. Bloch, J. Régis, et J. López-Krahe. From 3d magnetic resonance images to structural representations of the cortex topology using topology preserving deformations. *Journal of Mathematical Imaging and Vision*, 5(4) :297–318, 1995. (Cité page 44.)
- H. Markram. The blue brain project. *Nature Reviews Neuroscience*, 7(2) : 153–159, 2006. (Cité page 11.)
- L. Marrakchi-Kacem, F. Poupon, P. Roca, A. Tucholka, C. Delmaire, E. Bardinnet, M. Sharman, R. Valabregue, A. Messé, C. Malherbe, et al. Probabilistic atlas of the connections between the basal ganglia and the cortex. 2010. (Cité page 173.)
- M. Matelli, G. Luppino, et G. Rizzolatti. Architecture of superior and mesial area 6 and the adjacent cingulate cortex in the macaque monkey. *The Journal of comparative neurology*, 311(4) :445–462, 1991. ISSN 1096-9861. (Cité page 24.)
- M.M. Mesulam. Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of neurology*, 28(5) : 597–613, 1990. ISSN 1531-8249. (Cité page 13.)
- S. Mori, B.J. Crain, VP Chacko, et P. Van Zijl. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of neurology*, 45(2) :265–269, 1999. (Cité page 53.)
- V.B. Mountcastle. An organizing principle for cerebral function : the unit module and the distributed system. *The mindful brain*, pages 7–50, 1978. (Cité page 13.)
- V.B. Mountcastle. The columnar organization of the neocortex. *Brain*, 120 (4) :701, 1997. ISSN 0006-8950. (Cité page 37.)
- V.B. Mountcastle. *Perceptual neuroscience : The cerebral cortex*. Harvard Univ Pr, 1998. ISBN 0674661885. (Cité page 11.)
- L. Nanetti, L. Cerliani, V. Gazzola, R. Renken, et C. Keysers. Group analyses of connectivity-based cortical parcellation using repeated k-means clustering. *Neuroimage*, 47(4) :1666–1677, 2009. (Cité pages 23, 24, 184 et 185.)
- J. Neumann, D.Y. Von Cramon, B.U. Forstmann, S. Zysset, et G. Lohmann. The parcellation of cortical areas using replicator dynamics in fMRI. *Neuroimage*, 32(1) :208–219, 2006. ISSN 1053-8119. (Cité page 20.)

- Grégory Operto. *Analyse structurelle surfacique de données fonctionnelles cérébrales*. PhD thesis, 2008. (Cité pages 29 et 177.)
- C. Palm, M. Axer, D. Gräsel, J. Dammers, J. Lindemeyer, K. Zilles, U. Pietrzyk, et K. Amunts. Towards ultra-high resolution fibre tract mapping of the human brain—registration of polarised light images and reorientation of fibre vectors. *Frontiers in Human Neuroscience*, 4, 2010. (Cité page 178.)
- G.J.M. Parker et D.C. Alexander. Probabilistic anatomical connectivity derived from the microscopic persistent angular structure of cerebral tissue. *Philosophical Transactions of the Royal Society B : Biological Sciences*, 360(1457) :893, 2005. (Cité page 53.)
- Richard E. Passingham, Klaas E. Stephan, et Rolf Kötter. The anatomical basis of functional localization in the cortex. *Nature*, 3 :606–616, August 2002. (Cité page 23.)
- V. Perlberg et G. Marrelec. Contribution of exploratory methods to the investigation of extended large-scale brain networks in functional mri : Methodologies, results, and challenges. *International Journal of Biomedical Imaging*, 2008, 2008. Article ID 218519, 14 pages. (Cité page 26.)
- M. Perrin, Y. Cointepas, A. Cachia, C. Poupon, B. Thirion, D. Riviere, P. Cathier, V. El Kouby, A. Constantinesco, D. Le Bihan, et al. Connectivity-based parcellation of the cortical mantle using q-ball diffusion imaging. *Journal of Biomedical Imaging*, 2008 :1–10, 2008. ISSN 1687-4188. (Cité pages 23, 24, 44, 54, 55, 56, 185 et 186.)
- M. Perrin, C. Poupon, Y. Cointepas, B. Rieul, N. Golestani, C. Pallier, D. Riviere, A. Constantinesco, D. Le Bihan, et J.F. Mangin. Fiber tracking in q-ball fields using regularized particle trajectories. Dans *Information Processing in Medical Imaging*, pages 52–63. Springer, 2005. (Cité pages 53, 54 et 55.)
- Muriel Perrin. *Imagerie de diffusion à haute résolution angulaire : étude du modèle q-ball par couplage simulations - fantômes et application au suivi de fibres et à la parcellisation du cortex*. PhD thesis, Université Paris Sur, Faculté de sciences d’Orsay, France, 2006. (Cité pages 54 et 129.)
- N. Picard et P.L. Strick. Motor areas of the medial wall : a review of their location and functional activation. *Cerebral Cortex*, 6(3) :342, 1996. ISSN 1047-3211. (Cité pages 24 et 164.)
- C. Poupon. *Détection des faisceaux de fibres de la substance blanche pour l’étude de la connectivité anatomique cérébrale*. PhD thesis, Ecole Nationale Supérieure des Télécommunications, Paris , France, Dec. 1999. (Cité pages 42, 51 et 53.)
- C. Poupon, F. Poupon, L. Allirol, et J.-F. Mangin. A database dedicated to anatomo-functional study of human brain connectivity. Dans *12th HBM Neuroimage*, numéro 646, Florence, Italie, 2006. (Cité page 38.)

- M.E. Raichle, A.M. MacLeod, A.Z. Snyder, W.J. Powers, D.A. Gusnard, et G.L. Shulman. A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, 98(2) :676, 2001. (Cité page 20.)
- M.E. Raichle et A.Z. Snyder. A default mode of brain function : a brief history of an evolving idea. *Neuroimage*, 37(4) :1083–1090, 2007. (Cité page 20.)
- D. Rivière, J.F. Mangin, D. Papadopoulos-Orfanos, J.M. Martinez, V. Frouin, et J. Régis. Automatic recognition of cortical sulci of the human brain using a congregation of neural networks. *Medical Image Analysis*, 6(2) :77–92, 2002. ISSN 1361-8415. (Cité pages 19 et 44.)
- P. Roca, D. Riviere, P. Guevara, C. Poupon, et J.F. Mangin. Tractography-based parcellation of the cortex using a spatially-informed dimension reduction of the connectivity matrix. Dans *12thProc. MICCAI, LNCS Springer Verlag*, London, England, 2009a. (Cité pages 55, 129, 131 et 141.)
- P. Roca, D. Rivière, P. Guevara, C. Poupon, et J.-F. Mangin. Inter-subject connectivity-based parcellation of a patch of cerebral cortex. Dans *MICCAI 2010, LNCS 5762, Springer-Verlag*, 2010a. (Cité pages 55, 129, 145 et 154.)
- P. Roca, Denis Rivière, Pamela Guevara, Cyril Poupon, et J.-F. Mangin. Dwi-based parcellation of the human cortex with a new dimension reduction of the connectivity matrix. Dans *Proc. 15th HBM*, 2009b. (Cité pages 55 et 65.)
- P. Roca, Alan Tucholka, D. Riviere, P. Guevara, C. Poupon, et J.F. Mangin. Inter-subject connectivity-based parcellation of the human post-central gyrus. Dans *Proc. 16th HBM*, Barcelona, Spain, 2010b. (Cité page 55.)
- M. Rubinov et O. Sporns. Complex network measures of brain connectivity : uses and interpretations. *Neuroimage*, 2009. (Cité page 14.)
- M. F. S. Rushworth, T. E. J. Behrens, et Heidi Johansen-Berg. Connections patterns distinguish 3 regions of human parietal cortex. *Cerebral Cortex*, 16 :1418–1430, 2005. (Cité page 24.)
- SA Sarkisov, IN Filimonoff, et NS Preobrashenskaya. Cytoarchitecture of the human cortex cerebri. *Moscow : Medgiz*, 1949. (Cité page 19.)
- A. Schleicher, P. Morosan, K. Amunts, et K. Zilles. Quantitative architectural analysis : a new approach to cortical mapping. *Journal of autism and developmental disorders*, 39(11) :1568–1581, 2009. (Cité page 30.)
- Axel Schleicher, Katrin Amunts, Stefan Geyer, P. Morosan, et Karl Zilles. Observer-independent method for microstructural parcellation of cerebral cortex : A quantitative approach to cytoarchitectonics. *NeuroImage*, 9 :165–177, 1999. (Cité page 19.)
- R.I. Schubotz, A. Anwander, T.R. Knösche, D.Y. Von Cramon, et M. Tittgemeyer. Anatomical and functional parcellation of the human lateral

- premotor cortex. *NeuroImage*, 50(2) :396–408, 2010. ISSN 1053-8119. (Cité pages 23 et 184.)
- G.E. Smith. A new topographical survey of the human cerebral cortex, being an account of the distribution of the anatomically distinct cortical areas and their relationship to the cerebral sulci. *Journal of anatomy and physiology*, 41(Pt 4) :237, 1907. (Cité page 19.)
- S.M. Smith, C.F. Beckmann, N. Ramnani, M.W. Woolrich, P.R. Bannister, M. Jenkinson, P.M. Matthews, et D.J. McGonigle. Variability in fMRI : A re-examination of inter-session differences. *Human Brain Mapping*, 24(3) :248–257, 2005. ISSN 1097-0193. (Cité page 29.)
- O. Sporns, D.R. Chialvo, M. Kaiser, et C.C. Hilgetag. Organization, development and function of complex brain networks. *Trends in Cognitive Sciences*, 8(9) :418–425, 2004. ISSN 1364-6613. (Cité page 13.)
- O. Sporns et J.D. Zwi. The small world of the cerebral cortex. *Neuroinformatics*, 2(2) :145–162, 2004. (Cité pages 13, 14 et 77.)
- Olaf Sporns, Giulio Tononi, et Rolf Kötter. The human connectome : A structural description of the human brain. *PLOS Computational Biology*, 1(4) :245–251, September 2005. (Cité pages 9 et 10.)
- A. Steger et N.C. Wormald. Generating random regular graphs quickly. *Combinatorics, Probability and Computing*, 8(04) :377–396, 1999. ISSN 0963-5483. (Cité page 86.)
- EO Stejskal et JE Tanner. Spin diffusion measurements : spin echoes in the presence of a time-dependent field gradient. *The Journal of Chemical Physics*, 42(1) :288, 1965. ISSN 0021-9606. (Cité pages 42 et 43.)
- J. Talairach et P. Tournoux. Co-planar stereotaxic atlas of the human brain : 3-dimensional proportional system : an approach to cerebral imaging. 1988. (Cité page 19.)
- J. Tanji. Sequential organization of multiple movements : involvement of cortical motor areas. *Annual Review of Neuroscience*, 24(1) :631–651, 2001. ISSN 0147-006X. (Cité page 24.)
- B. Thirion, P. Pinel, S. Mériaux, A. Roche, S. Dehaene, et J.B. Poline. Analysis of a large fMRI cohort : Statistical and methodological issues for group analyses. *Neuroimage*, 35(1) :105–120, 2007. ISSN 1053-8119. (Cité page 29.)
- Bertrand Thirion, Guillaume Flandin, Philippe Pinel, Alexis Roche, Philippe Ciuciu, et Jean-Baptiste Poline. Dealing with the shortcomings of spatial normalization : multi-subject parcellation of fmri datasets. *Human Brain Mapping*, 27 :678–693, 2006. (Cité pages 20 et 177.)
- V. Tomassini, S. Jbabdi, J.C. Klein, T.E.J. Behrens, C. Pozzilli, P.M. Matthews, M.F.S. Rushworth, et H. Johansen-Berg. Diffusion-weighted imaging tractography-based parcellation of the human lateral premotor cortex identifies dorsal and ventral subregions with anatomical and functional specializations. *Journal of Neuroscience*, 27(38) :10259, 2007. (Cité pages 23 et 184.)

- David Tuch. *Diffusion MRI of complex tissue structure*. PhD thesis, Division of Health Sciences and Technology, University of Chicago, Massachusetts INstitute of Technology, 2002. (Cité pages 51 et 52.)
- D.S. Tuch. Q-ball imaging. *Magnetic Resonance in Medicine*, 52(6) :1358–1372, 2004. (Cité page 52.)
- D.S. Tuch, T.G. Reese, M.R. Wiegell, et W. Van J. Diffusion mri of complex neural architecture. *Neuron*, 40(5) :885–895, 2003. (Cité page 52.)
- A. Tucholka, B. Thirion, M. Perrot, P. Pinel, J.-F. Mangin, et J.-B. Poline. Probabilistic anatomo-functional parcellation of the cortex : how many regions? Dans *11thProc. MICCAI, LNCS Springer Verlag*, New-York, USA, 2008. (Cité pages 20 et 22.)
- Alan Tucholka. *Prise en compte de l'anatomie cérébrale individuelle dans les études d'IRM fonctionnelle*. PhD thesis, Université Paris-Sud 11, Faculté des Sciences d'Orsay, 2010. (Cité pages 21, 38, 45 et 46.)
- N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, et M. Joliot. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*, 15(1) :273–289, 2002. (Cité page 18.)
- P.A. Valdés-Sosa, J.M. Sánchez-Bornot, A. Lage-Castellanos, M. Vega-Hernández, J. Bosch-Bayard, L. Melie-García, et E. Canales-Rodríguez. Estimating brain functional connectivity with sparse multivariate autoregression. *Philosophical Transactions B*, 360(1457) :969, 2005. (Cité page 77.)
- David C. Van Essen et Donna L. Dierker. Surface-based and probabilistic atlases of primate cerebral cortex. *Neuron*, 56 :209–225, October 2007. (Cité page 133.)
- D.C. Van Essen. A tension-based theory of morphogenesis and compact wiring in the central nervous system. *NATURE-LONDON*-, pages 313–318, 1997. ISSN 0028-0836. (Cité page 28.)
- W.N. Venables et B.D. Ripley. *Modern applied statistics with S*. Springer verlag, 2002. ISBN 0387954570. (Cité page 111.)
- C. Vogt et O. Vogt. *Allgemeine ergebnisse unserer hirnforschung*. 1919. (Cité pages 19 et 24.)
- V. Vorobiev, P. Govoni, G. Rizzolatti, M. Matelli, et G. Luppino. Parcellation of human mesial area 6 : cytoarchitectonic evidence for three separate areas. *European Journal of Neuroscience*, 10(6) :2199–2203, 1998. ISSN 1460-9568. (Cité page 24.)
- Mark T. Wallace. A revised view of sensory cortical parcellation. *PNAS*, 101 :2167–2172, February 2004. (Cité page 13.)

- N.B. Walters, S.B. Eickhoff, A. Schleicher, K. Zilles, K. Amunts, G.F. Egan, et J.D.G. Watson. Observer-independent analysis of high-resolution MR images of the human cerebral cortex : in vivo delineation of cortical areas. *Human brain mapping*, 28(1) :1–8, 2007. ISSN 1097-0193. (Cité page 19.)
- D.J. Watts et S.H. Strogatz. Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684) :440–442, 1998. (Cité page 15.)
- C. Wernicke. The aphasic symptom-complex : a psychological study on an anatomical basis. *Archives of Neurology*, 22(3) :280, 1970. (Cité pages 13, 14 et 20.)
- P.I. Yakovlev et A.R. Lecours. The myelogenetic cycles of regional maturation of the brain. *Regional development of the brain in early life*, pages 3–70, 1967. (Cité page 31.)
- A. Zalesky, A. Fornito, I.H. Harding, L. Cocchi, M. Yücel, C. Pantelis, et E.T. Bullmore. Whole-brain anatomical networks : Does the choice of nodes matter ? *Neuroimage*, 2010. (Cité page 18.)
- K. Zilles et K. Amunts. Receptor mapping : architecture of the human cerebral cortex. *Current opinion in neurology*, 22(4) :331, 2009. ISSN 1350-7540. (Cité page 19.)
- K. Zilles, G. Schlaug, S. Geyer, G. Luppino, M. Matelli, M. Qü, A. Schleicher, et T. Schormann. Anatomy and transmitter receptors of the supplementary motor areas in the human and nonhuman primate brain. *Advances in neurology*, 70 :29, 1996. ISSN 0091-3952. (Cité page 24.)
- Karl Zilles et Katrin Amunts. Centenary of brodmann’s map — conception and fate. *Nature Reviews Neuroscience*, 2010. AOP, published online 4 January 2010. (Cité page 29.)

PUBLICATIONS

CONFÉRENCES AVEC COMITÉ DE LECTURE

P. Roca, D. Rivière, P. Guevara, C. Poupon, ans J.-F. Mangin. Tractography-Based Parcellation of the Cortex using a Spatially-Informed Dimension Reduction of the Connectivity Matrix. *In Proc. 12th International Conference on Medical Image Computing and Computer Assisted Intervention*, 8 pages, *Lecture Notes in Computer Science*, Springer Verlag, London, England, 2009.

P. Roca, A. Tucholka, D. Rivière, P. Guevara, C. Poupon, ans J.-F. Mangin. Inter-subject Connectivity-based Parcellation of a Patch of Cerebral Cortex. *In Proc. 13th International Conference on Medical Image Computing and Computer Assisted Intervention*, Beijing, 8 pages, *Lecture Notes in Computer Science*, Springer Verlag, China, 2010.

L. Marrakchi-Kacem, C. Delmaire, A. Tucholka, P. Roca, P. Guevara, F. Poupon, J. Yelnik, A. Durr, J.F. Mangin, S. Lehericy, et al. *Analysis of the striato-thalamo-cortical connectivity on the cortical surface to infer biomarkers of huntington's disease. Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010*, pages 217–224, 2010.

CONFÉRENCES SANS COMITÉ DE LECTURE

P. Roca, D. Rivière, P. Guevara, C. Poupon, ans J.-F. Mangin. DWI-based parcellation of the human cortex with a new dimension reduction of the connectivity matrix. *In proc. 15th Organization for Human Brain Mapping*, San Francisco, California, 2009.

P. Roca, A. Tucholka , D. Rivière, P. Guevara, C. Poupon, ans J.-F. Mangin. Inter-subject Connectivity-based Parcellation of the Post-central Gyrus. *In proc. 16th Organization for Human Brain Mapping*, Barcelona, Spain, 2010.

L. Marrakchi-Kacem, F. Poupon, A. Tucholka, P. Roca, C. Delmaire, J. Yelnik, E. Bardinet, Al. Durr, J.-F. Mangin, S. Lehericy, C. Poupon. Striato-Cortical Probabilistic Connectivity Atlases to infer Biomarkers of Huntington's Disease. *In Proc. 16th Organization for Human Brain Mapping*, Barcelona, Spain, 2010.

L. Marrakchi-Kacem, F. Poupon, A. Tucholka, P. Roca, C. Delmaire, J. Yelnik, E. Bardinnet, Al. Durr, J.-F. Mangin, S. Lehericy, C. Poupon. Probabilistic Atlas of the Connections between the basal ganglia and the cortex. *In Proc. 18th International Society for Magnetic Resonance in Medicine (ISMRM-ESMRMB Joint Annual Meeting), Stockholm, Sweden, 2010.*

PARTICIPATION À DES ARTICLES DE JOURNAUX

7.E. Jouvent, S. Reyes, J.-F. Mangin, P. Roca, M. Perrot, B. Thyreau, D. Hervé, M. Dichgans, H. Chabriat. Apathy Is Related To Cortex Morphology In CADASIL : A Sulcal -Based Morphometry Study. *Neurology*, 76(17) :1472, 2011.

NOTATIONS ET ABRÉVIATIONS

NOTATIONS

$\mathbf{R}_{\text{SMA} \cup \text{pré-SMA}}$	Région englobant le SMA et le pré-SMA
$S = (\mathcal{V}, \mathcal{E}_{\text{spatial}})$	Surface du cortex
$\mathcal{V} = (v_i)_{i=1..V}$	points ou noeuds du maillage
$\mathcal{E}_{\text{spatial}}$	relations spatiales entre les noeuds du maillage (polygones)
$T = [p_i]_{i=1..I}$	Tract : Liste de points $p_i \in \mathbb{R}^3$
T_1, \dots, T_N	Faisceau : ensemble de N tracts
$p_{v_i} \in (\mathbb{N}^+)^V$	Profil de connectivité cortical d'un point $v_i \in \mathcal{V}$
	$p_{v_i}(j)$ = nombre de tracts reliant les points v_i et v_j
$C_{\mathcal{V}} \in M_{(V,V)}(\mathbb{N})$	Matrice de connectivité corticale (entre les points v_i)
	$C_{\mathcal{V}}(i, j)$ = nombre de tracts reliant v_i et v_j
	$i^{\text{ème}}$ ligne de $C_{\mathcal{V}} = p_{v_i}$
$C_{\mathcal{V}, \text{bin}} \in M_{(V,V)}(\{0, 1\})$	Matrice des connexions corticales (binarisation de $C_{\mathcal{V}}$)
$C_{\mathcal{V}, \text{density}}$	Matrice des densités de connexions corticale
$\mathcal{A} = \cup_{p=1..P} A_p$	Partition de \mathcal{V} en P parcelles appelée Parcellisation
$(\text{surf}_A)_{A \in \mathcal{A}}$	Aires des parcelles de la partition \mathcal{A}
$C_{\mathcal{A}} \in M_{(P,P)}(\mathbb{N})$	Matrice de connectivité entre les parcelles
	$C_{\mathcal{A}}(p, q)$ = nombre de tracts reliant A_p et A_q

ABRÉVIATIONS

PGSE	séquence Pulse Gradient Spin Echo
SMA	Aire motrice supplémentaire
DSI	Diffusion Spectrum Imaging (Imagerie du Spectre de Diffusion)
ODF	Orientation Distribution Function (Fonction de Distribution des Orientations des fibres)
d-ODF	Diffusion Orientation Distribution Function (Fonction de Distribution des Orientations de Diffusion)
f-ODF	Fiber Orientation Distribution Function (Fonction de Distribution des Orientations de Fibres)
DTI	Diffusion Tensor Imaging (Imagerie du Tenseur de Diffusion)
HARDI	High Angular Resolution Diffusion Imaging (Imagerie de Diffusion à Haute Résolution Angulaire)
HS	Harmoniques Sphériques
LPE	Ligne de partage des eaux
ACP	Analyse en composantes principales
R	Language et logiciel libre spécialisé en statistique.

LISTE DES FIGURES

1.1	Techniques de construction du connectome cérébral humain	9
1.2	Les différentes parties de l'encéphale	10
1.3	Les trois milieux constituant le cerveau	11
1.4	Les échelles spatiales du connectome	12
1.5	Construction d'un connectome de 66 ou 998 noeuds	16
1.6	Distribution des degrés et forces de connexion	17
1.7	Répartition spatiale des degrés et forces de connexion	17
1.8	Echelle micro et macroscopique	21
1.9	Caractérisation fonctionnelle des aires de Brodmann	22
1.10	Parcellisation fonctionnelle de Broca	22
1.11	Région englobant le SMA et le pré-SMA	25
1.12	Etude 1. Parcellisation et connexions du cortex frontal médial	26
1.13	Etude 2. Parcellisation du cortex frontal médial	27
1.14	Etude 2. Connexions des régions du cortex frontal médial .	27
1.15	Variabilité sulco-gyrale interindividuelle	28
1.16	De multiples parcellisations cytoarchitectoniques du cortex	29
1.17	Exemple des régions couvertes par un atlas probabiliste . .	31
1.18	Cartes de probabilité surfaciques d'aires de Brodmann . . .	31
1.19	Connexions anatomiques et variabilité sur histologie/DTI .	32
1.20	Variabilité inter-sujets des connexions de l'aire de Broca . .	33
2.1	Surface corticale et tracts	39
2.2	Relaxation des spins	40
2.3	Temps de relaxation des tissus cérébraux	41
2.4	Matière blanche et diffusion anisotrope	42
2.5	Séquence Pulse Gradient Spin Echo (PGSE)	43
2.6	Segmentation du cerveau de Brainvisa	47
2.7	Modélisation inter-sujets surfacique de Freesurfer	48
2.8	Région d'intérêt englobant le SMA et le pré-SMA	49
2.9	Différents types de données locales de diffusion	50
2.10	Modèle du tenseur de diffusion	51
2.11	Tractographie par ligne de courant	54
2.12	Récapitulatif des prétraitements	56
3.1	Vocabulaire. Parcelle, profil et matrice de connectivité	64
3.2	Diagramme de Voronoï centroïdal	67
3.3	Calcul de la matrice de connectivité corticale	69
3.4	Représentation de ConnectomeViewer)	70
3.5	Autres visualisations 2D du connectome	70
3.6	Atlas macroscopique du connectome - Visualisation	72

4.1	Données simulées d'un sujet	78
4.2	Principaux objets simulés. Notations	79
4.3	Les différentes étapes de la simulation d'un sujet.	83
4.4	Variabilité inter-sujets de la simulation	85
4.5	Choix et variation des paramètres des données simulées. . .	88
4.6	Deux bases simulées : cas optimal et extrême	89
5.1	Exemple de patches corticaux	94
5.2	Exemple de matrice de connectivité d'un patch	94
5.3	Réduction de dimension intra-sujet par identification de ré- gions d'intérêt	95
5.4	Principe de la ligne de partage des eaux (LPE)	96
5.5	Bassins LPE et tracts associés	96
5.6	Réduction de dimension inter-sujets par identification de ré- gions d'intérêt	98
5.7	Analyse d'un gyrus de la base simulée	99
5.8	Avant/après réduction de dimension. Clusplots	100
5.9	Avant/après réduction de dimension. Silhouettes.	101
5.10	Résultats inter-sujets. Base simulée optimale	102
5.11	Base optimale. ACP inter-sujets (Clusplot)	102
5.12	Base optimale. Réduction LPE intra et inter-sujets	103
5.13	Base extrême. Vraies parcellisations	104
5.14	Base extrême. Profils de connectivité du gyrus 1	104
5.15	Base extrême. Bassins LPE intra- et inter-sujets	105
5.16	Bas extrême. PCA intra-sujet (Clusplot)	105
5.17	Gyrus post-central. Profils de connectivité et bassins LPE inter-sujets	106
5.18	Intérêt de la LPE inter-sujets	108
5.19	SMA et pré-SMA. Profils et bassins LPE intra- et inter-sujets	109
5.20	SMA et pré-SMA. Profils réduits par LPE	110
5.21	Diagramme de Voronoï centroïdal sur un hémisphère	112
5.22	Parcellisation simulée de 200 régions par Voronoï Centroïdal	114
5.23	Base optimale. ACP intra-sujet. Composantes et variance des données.	114
5.24	Base optimale. ACP intra-sujet. Composantes et connexions	115
5.25	Base optimale. ACP intra-sujet. Clusplot.	115
5.26	Base optimale. ACP inter-sujets. Clusplot.	116
5.27	Base optimale. ACP inter-sujets. Composantes et variance des données.	117
5.28	Base extrême. Composantes et variance des données. . . .	118
5.29	Base extrême. PCA intra-sujet. Clusplot.	118
5.30	Bas extrême. PCA inter-sujets. Clusplot.	119
5.31	Gyrus post-central. ACP intra-sujet. Composantes et va- riance des données.	120
5.32	Gyrus post-central. ACP intra-sujet. Composantes principales.	121
5.33	Gyrus post-central. Profil de connectivité.	121
5.34	Gyrus post-central. ACP intra-sujet. Scores des composantes.	122
5.35	Distance de Hausdorff	124
5.36	Comparaison distance eculidienne/distance de Hausdorff. .	125
5.37	Base optimale et distances entre profils.	126

5.38	Base extrême. Cross-corrélation entre les profils pour différentes méthodes de réduction.	127
5.39	Profils et parcellisations.	129
5.40	Différentes approches de parcellisation.	130
6.1	Parcellisation itérative intra-sujet. Organigramme.	134
6.2	Parcellisation intra-sujet sur la base simulée optimale. . . .	140
6.3	Parcellisation intra-sujet sur la base simulée extrême. . . .	140
6.4	Histogramme des longueurs des tracts.	141
6.5	Parcellisation intra-sujet. Résultat pour un sujet.	142
6.6	Parcellisation intra-sujet : Sensibilité à l'initialisation. . . .	142
6.7	Parcellisation intra-sujet. Résultats de 4 sujet autour de Broca. .	143
7.1	Réduction de dimension et clustering inter-sujets	149
7.2	Différentes combinaisons des profils d'un groupe de sujets. .	150
7.3	Base simulée optimale. Index de Rand des parcellisations inter-sujets.	152
7.4	Base simulée optimale. Parcellisation inter-sujets vs gold standard d'un gyrus.	152
7.5	Base simulée extrême. Parcellisations inter-sujets.	154
7.6	Parcellisation inter-sujets du gyrus post-central par concaténation. Parcelles obtenues.	156
7.7	Gyrus post-central. Connexions des parcelles inter-sujets et profils concaténés réordonnés.	157
7.8	Parcellisation inter-sujets du gyrus post-central par moyennage.	157
7.9	Une méthode de parcellisation pour un type de réduction des profils.	158
7.10	Comparaison des méthodes inter-sujets selon la densité des connexions.	160
7.11	Gyrus 1. Parcellisations des différentes méthodes pour une densité de connexion.	160
7.12	Méthode M_{ROIs} sur 10 sujets par gyrus selon la densité de connexions.	161
7.13	Méthode M_{ROIs} sur 20 sujets par gyrus selon la densité de connexions.	162
7.14	Comparaison des méthodes avec différents niveau de variabilité inter-individuelle.	163
7.15	Parcellisations inter-sujets de $\mathbf{R}_{SMAUp\acute{r}\acute{e}-SMA}$	166
7.16	Parcellisations inter-sujets de $\mathbf{R}_{SMAUp\acute{r}\acute{e}-SMA}$. Profils de connectivité.	167
7.17	ACP sur les profils de $\mathbf{R}_{SMAUp\acute{r}\acute{e}-SMA}$ pour un sujet. Composantes.	168
7.18	ACP sur les profils de $\mathbf{R}_{SMAUp\acute{r}\acute{e}-SMA}$ pour un sujet. Scores. .	169
7.19	Clustering intra-sujet après PCA de $\mathbf{R}_{SMAUp\acute{r}\acute{e}-SMA}$	169

Ce document a été préparé à l'aide de l'éditeur de texte Kate et du logiciel de composition typographique L^AT_EX 2_ε.

Titre Parcellisation du manteau cortical à partir de la connectivité anatomique cartographiée par imagerie de diffusion

Résumé La parcellisation du cerveau humain en aires fonctionnelles est un problème complexe mais majeur pour la compréhension du fonctionnement du cerveau et pourrait avoir des applications médicales importantes en neurochirurgie par exemple pour mieux identifier les zones fonctionnelles à sauvegarder. Cet objectif va de pair avec l'idée de construire le connectome cérébral humain, qui n'est autre que le réseau de ses connexions. Pour définir un tel réseau, il faut en effet définir les éléments de ce réseau de connexions : c'est-à-dire avoir un découpage du cerveau en régions. Il existe de multiples manières et critères pour identifier ces régions et à ce jour il n'y a pas de parcellisation universelle du cortex. Dans cette thèse nous étudierons la possibilité d'effectuer cette parcellisation en fonction des données de connectivité anatomique, issues de l'imagerie par résonance magnétique de diffusion, qui est une technique d'acquisition permettant de reconstruire les faisceaux de neurones cérébraux de manière non invasive. Nous nous placerons dans un cadre surfacique en étudiant seulement la surface corticale et les connexions anatomiques sous-jacentes. Dans ce contexte nous présenterons un ensemble de nouveaux outils pour construire, visualiser et simuler le connectome cérébral humain, dans un cadre surfacique et à partir des données de connectivité anatomique reconstruites par IRM, et ceci pour un groupe de sujets. A partir de ces outils nous présenterons des méthodes de réduction de dimension des données de connectivité, que nous appliquerons pour parcelliser le cortex entier de quelques sujets. Nous proposons aussi une nouvelle manière de décomposer les données de connectivité au niveau d'un groupe de sujets en tenant compte de la variabilité inter-individuelle. Cette méthode sera testée et comparée à d'autres méthodes sur des données simulées et des données réelles. Les enjeux de ce travail sont multiples, tant au niveau méthodologique (comparaison de différents algorithmes de tractographie par exemple) que clinique (étude du lien entre altérations des connexions et pathologie).

Mots-clés parcellisation corticale, IRM de diffusion, connectome cérébral humain, réduction de dimension, profils de connectivité

Title Connectivity-based parcellation of the human cortex

Abstract In-vivo parcellation of the human cortex into functional brain areas is a major goal to better understand how the brain works and could have a lot of medical applications and give useful information to guide neurosurgery for example. This objective is related to the building of the human brain connectome, which is the networks of brain connections. Indeed, it is necessary to define the basic element of this connectome, and for doing this to have a subdivision of the cortex into brain regions. Actually, there is no such gold standard parcellation : there are a lot of techniques and methods to achieve this goal. During this PhD., anatomical connectivity based on diffusion-weighted imaging has been used to address this problem, with a surfacic approach. In this context, we will present a set of new tools to create, visualize and simulate the human brain connectome for a group of subjects. We will introduce dimension reduction methods to compile the cortical connectivity profiles taking into account the interindividual variability. These methods will be applied to parcellate the cortex, for one subject or for a group of subjects simultaneously. There are many applications of this work, in methodology, to compare tractography algorithms for example or in clinical, to look at the relations between connections damages and pathology.

Keywords cortical parcellation, diffusion MRI, human brain connectome, dimension reduction, connectivity profiles